PRICING AND REVENUE MANAGEMENT IN SUPPLY CHAIN NETWORKS AND SERVICE SYSTEMS

A Dissertation
Presented to
The Academic Faculty

By

Pornpawee Bumpensanti

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Industrial and Systems Engineering

Georgia Institute of Technology

May 2021

PRICING AND REVENUE MANAGEMENT IN SUPPLY CHAIN NETWORKS AND SERVICE SYSTEMS

Thesis committee:

Dr. He Wang, Advisor
School of Industrial and Systems Engineering
Georgia Institute of Technology

Dr. Martin Savelsbergh
School of Industrial and Systems Engineering
Georgia Institute of Technology

Dr. Pinar Keskinocak
School of Industrial and Systems Engineering
Georgia Institute of Technology

Dr. Cong Shi
Department of Industrial and Operations Engineering
University of Michigan

Dr. Anton Kleywegt
School of Industrial and Systems Engineering
Georgia Institute of Technology

Date approved: April 29, 2021

This thesis is dedicated to my mom and dad,

who always support whatever journey I choose to pursue.

ACKNOWLEDGMENTS

First and foremost, I would like to extend my deepest appreciation to my advisor, He Wang, for all the guidance and continuing support during my PhD journey. In the past five years, he has taught me so many things that help me grow both personally and professionally. His enthusiasm and motivation have always been my great inspiration. It has been an honor working with you.

I would also like to extend my sincere thanks to Pinar Keskinocak, Anton Kleywegt, Martin Savelsbergh and Cong Shi for serving on my thesis committee. I really appreciate their valuable time going through this thesis and providing helpful feedback. Their constructive comments and suggestions really made this thesis so much richer. Additionally, thanks to the ISyE faculty for providing me with solid foundations for my research. I am very fortunate to learn many interesting courses from the pioneers of their fields.

My gratitude goes to my co-authors: Martin Savelsbergh, Siva Theja Maguluri and Sushil Mahavir Varma. They are indeed outstanding researchers. These research collaborations have been pleasant and valuable experiences to me.

Life away from home is not easy, but it is not so difficult, because I am surrounded by great friends in Atlanta. I am very thankful for their support and encouragement. Special thanks to Yuwei Zhou and Daniela Hurtado Lange for patiently listening to me complaining about life and cheering me up. I would also like to thank Thai student community at Georgia Tech for making me feel at home in Atlanta. Our Friday dinner always helps refresh my mind. Life here without Wundzen Juthamas Kunahirunkanok, Max Chayanont Pithaksithiporn, Sek Sangratkanjanasin, Dhorn Kosolpatanadurong and Gio Weerapich Panlertkitsakul would have been much more boring. Thank you for constantly and frequently hanging out with me. Moreover, I would like to thank Ep Jintasit Pravitra for helping me and everyone here whenever needed. Special thanks to Nat Natthapol Prakongpan for taking a very good care of me.

I also have so many amazing friends from Thailand. Thank you for always checking up on me and giving me words of encouragement. I would especially like to extend my warm thanks to Neoi Sittiphol Phanvilai and Wien Wiennat Mongkulmann for their meaningful emotional support and help with my deficiency in technology skills. Also, I would like to thank Pherm Phermsak Lilakul who always helps with my English. My special appreciation goes to Thaisiri Watewai for helping me academically and non-academically.

I am forever indebted to my dear family. Thanks to my brother, Noy Palintorn Bumpensanti, for his caring. Most importantly, I would like to extend my deepest gratitude to my parents, Suttinee Bumpensanti and Sanpong Bumpensanti, for their unconditional love, continuing first-class support and everything. Without you, this PhD would not have been possible.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

This thesis comprises four topics focusing on pricing and revenue management, a powerful technique to enhance firm profitability through demand and supply management. Because of its efficacy, it has been adopted by small and big companies in a broad range of industries.

In the first part of the thesis, joint work with He Wang, we consider a canonical quantity-based network revenue management problem, where a firm accepts or rejects incoming customer requests irrevocably in order to maximize expected revenue given limited resources. We design a new heuristic, which builds upon a family of re-solving heuristics that periodically re-optimize a deterministic approximation to the original problem. Our heuristic proves to have a strong theoretical performance guarantee and desirable numerical results.

The second part is joint work with Martin Savelsbergh and He Wang. We consider an integrated pricing and routing problem on a network which is motivated by applications in freight transportation. We propose two algorithms for the solution of this problem: a Frank-Wolfe type algorithm, and a primal-dual algorithm using an online learning technique. Both algorithms prove to have desirable convergence rates. Numerical experiments show significant profit improvement of integrated pricing and routing decisions over independent pricing or routing strategies.

In the third part of this thesis, joint work with Sushil Mahavir Varma, Siva Theja Maguluri and He Wang, we study a two-sided queueing system under joint pricing and matching controls. This problem is motivated by applications in the gig economy and online marketplaces. We propose a two-price and max-weight matching policy which proves to achieve the optimal rate. The proposed algorithm also shows promising numerical results when compared to other algorithms in similar settings.

The fourth part of the thesis is joint work with He Wang. We study a discrete choice model. The problem is motivated by the violation of regularity property, which states

that adding an option to a choice set cannot increase the choice probability for any of the original choice options. This property can be observed in real-world applications but cannot be represented by a widely-used random utility maximization (RUM) model. We propose a more general choice model – a model based on a neural network framework – which can explain any choice phenomenon. Numerical experiments show the model consistently outperforms other models for both synthetic and real datasets.

CHAPTER 1
INTRODUCTION

Revenue management is concerned with allocating limited resources to meet demand of customers with the objective of maximizing the revenue. Traditional revenue management method generally assumes that resource capacity or supply are exogenously given. It typically deals with demand management through pricing or capacity allocation decisions. However, focusing only on demand side only helps increase revenue which is one part of profit equation. The other component of profit equation which is a cost reduction should also be emphasized. An operating cost reduction can be obtained from supply chain management which is concerned with the problem of sourcing and delivering resources to satisfy demand with the objective of minimizing cost. That is, revenue management and supply chain management interactively play an important role in increasing a firm's profit and should be simultaneously considered to create competitive advantages.

The main part of this thesis focuses on applying revenue management to supply chain, logistics networks and on-demand services ranging from traditional service industries to disruptive transportation network company. Chapter 2 concentrates on the model for traditional applications of revenue management such as airline, retail, advertising and hospitality industries. Chapter 3 revolves around pricing in delivery services industry. Chapter 4 focuses on on-demand services and ride-sharing platforms. We consider different models based on characteristic of each application. In airline, retail, advertising and hospitality industries, firm is only able to manage limited resources by price or capacity allocation decisions. Different form a traditional revenue management framework, delivery service company also has a flexibility to determine distribution decision, along with demand management decision, to enhance business profitability. However, resource capacities are no longer limited for ride-sharing platform as firm can set price to induce supply to the system.

Therefore, to increase the profit firm must determine jointly pricing decision of demand side and supply side as well as how to efficiently match such demand and supply.

A common feature of the main part of the thesis is independent demand model assumption. Independent demand model assumes that the demand of any product is irrelevant of the availability of other products. That is, it assumes that customers purchase a particular product and never substitute one product for another. However, in reality customers may opt to substitute between similar products, which are available. Therefore, this assumption is somewhat unrealistic and can potentially cause revenue loss. Therefore, it is important to understand customer behavior to successfully tackle revenue management problems. The last part of this thesis (Chapter 5) steps away from independent demand model and focuses on discrete choice model.

1.1 Overview of results

The first part of the thesis (Chapter 2), we consider a canonical quantity-based network revenue management problem, where a firm accepts or rejects incoming customer requests irrevocably in order to maximize expected revenue given limited resources. Due to the curse of dimensionality, the exact solution to this problem by dynamic programming is intractable when the number of resources is large. We study a family of re-solving heuristics that periodically re-optimize an approximation to the original problem known as the deterministic linear program (DLP), where random customer arrivals are replaced by their expectations. We find that, in general, frequently re-solving the DLP produces the same order of revenue loss as one would get without re-solving, which scales as the square root of the time horizon length and resource capacities. By re-solving the DLP at a few selected points in time and applying thresholds to the customer acceptance probabilities, we design a new re-solving heuristic whose revenue loss is uniformly bounded by a constant that is independent of the time horizon and resource capacities.

An integrated pricing and routing problem on a network is considered in the second part of the thesis (Chapter 3). The problem is motivated by applications in freight transportation such as package delivery and less-than-truckload shipping services. The decision maker sets a price for each origin-destination pair of the network, which determines the demand flow that needs to be served. The flows are then routed through the network given fixed arc capacities and costs. Demand for the same origin-destination pair can be routed along multiple paths in the network if desirable. The objective is to maximize the revenues from serving demand minus the transportation costs incurred given the capacity constraints. We propose two algorithms for the solution of this problem: (1) a Frank-Wolfe type algorithm, which requires the objective function to be smooth, and (2) a primal-dual algorithm using an online learning technique, which allows non-smooth objective functions. We prove that the first algorithm has a convergence rate of $O(1/T)$ and the second algorithm has a convergence rate of $O(\log T/T)$, where T is the number of iterations. Numerical experiments on randomly generated instances show that coordinating pricing and routing decisions can improve profits significantly compared to independent pricing or routing strategies.

Motivated by applications in gig economy and online marketplaces, we study a two-sided queueing system under joint pricing and matching controls in the third part of the thesis (Chapter 4). The queueing system is modeled by a bipartite graph, where the vertices represent customer or server types and the edges represent compatible customer-server matches. Customers and servers sequentially arrive to the system and enter separate queues according to their types. The arrival rates of different types depend on the prices set by the system operator and the expected waiting time. At any point in time, the system operator can choose certain customers to match with compatible servers. The objective is to maximize the long-run average profit for the system. We first propose a fluid approximation based pricing and max-weight matching policy, which achieves an $O(\sqrt{\eta})$ optimality rate when all the arrival rates are scaled by η. We further show that a two-price and max-weight matching policy achieves an improved $O(\eta^{1/3})$ optimality rate. Under a broad class of pricing policies, we prove that any matching policy has an optimality rate that is lower bounded by $\Omega(\eta^{1/3})$. Thus, the two-price and max-weight matching policy achieves the optimal rate with respect to η. We also demonstrate the advantage of max-weight matching over a randomized matching policy. Under the complete resource pooling condition, we show that max-weight matching achieves $O(\sqrt{n})$ and $O(n^{1/3})$ optimality rates for static and two-price policies, respectively, where n is the number of customer and server types. In comparison, the randomized matching policy may have an $\Omega(n)$ optimality rate.

A neural network choice model is considered in the last part of this thesis (Chapter 5). We study a discrete choice problem, where a customer chooses an option from a finite set of alternatives. The majority of discrete choice models are derived from the random utility maximization (RUM) principle. This assumption implies regularity property, i.e., adding an option to a choice set cannot increase the choice probability for any of the original choice options. However, many empirical evidences suggest that the regularity property may be violated. There are also several efforts to model choice behavior beyond the RUM class, but their predicting performances generally tie with the nature of dataset. We propose a more general choice model, which can explain any choice phenomenon. The model is based upon a neural network framework. Our numerical results show that using the proposed neural network choice model consistently outperforms other choice models, either RUM or non-RUM models, in both synthetic and real datasets.

CHAPTER 2

RE-SOLVING HEURISTIC WITH UNIFORMLY BOUNDED LOSS IN NETWORK REVENUE MANAGEMENT

2.1 Introduction

The *network revenue management* (NRM) problem (Williamson, 1992; Gallego and Ryzin, 1997) is a classical model that has been extensively studied in the revenue management literature for over two decades. The problem is concerned with maximizing revenue given limited resource and time, and has a wide range of applications in the airline, retail, advertising, and hospitality industries (see examples in Talluri and Van Ryzin, 2004). However, the exact solution to the NRM problem is difficult to compute when the number of resources is large. Heuristics proposed in the previous literature typically have optimality gaps, i.e., expected revenue losses compared to the optimal solution, that increase with the time horizon and the resource capacities. In this chapter, we propose a new heuristic for the NRM problem for which the revenue loss is independent of the time horizon and the resource capacities.

The NRM problem is stated as follows: there is a set of resources with finite capacities that are available for a finite time horizon. Heterogeneous customers arrive sequentially over time. Customers are divided into different classes based on their consumption of resources and the prices they pay. Each class of customer may request multiple types of resources and multiple units of each resource. Upon a customer's arrival, a decision maker must irrevocably accept or reject the customer. If the customer is accepted and there is enough remaining capacities, she consumes the resources requested and pays a fixed price associated with her class. Otherwise, if the customer is rejected, no revenue is collected and no resources are used. Unused resources at the end of the finite horizon are perishable and have no salvage value. The decision maker's objective is to maximize the expected revenue earned during the finite horizon.

We note that the formulation stated above is more specifically known as the "quantity-based" NRM problem. In another formulation referred to as the "price-based" NRM problem, the decision maker chooses posted prices rather than accept/reject decisions. The two formulations are different, but are equivalent in some special cases (Maglaras and Meissner, 2006). We focus on the quantity-based formulation in this chapter.

A classical application of the NRM problem is in airline seat revenue management (Williamson, 1992; Gallego and Ryzin, 1997). Here, the resources correspond to flight legs and the capacity corresponds to the number of seats on each flight. The resources are perishable on the date of flight departure. Arriving customers are divided into separate classes defined by combinations of itinerary and fare. A simple flight network of two flight legs and three itineraries is shown in Figure 2.1. The objective of the airline is to maximize the expected revenue earned from allocating available seats to different classes of customers. Notice that the problem cannot be decomposed for each individual flight leg, since some itineraries use multiple resources simultaneously (e.g., in Figure 2.1, customers traveling from A to C would request itinerary $A \to B \to C$). In practice, the huge size of airline networks makes solving this problem challenging.

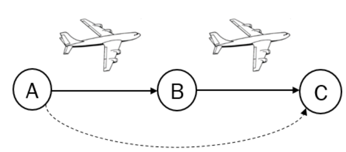


Figure 2.1: A flight network of two flight legs ($A \to B, B \to C$) and three itineraries.

2.1.1 Deterministic LP approximation and re-solving heuristics

In theory, the NRM problem can be solved by dynamic programming; however, since the state space grows exponentially with the number of resources, the dynamic programming formulation is often intractable. Therefore, we focus on heuristics with provable performance guarantees in this chapter. We define *revenue loss* as the gap between the expected revenue of a heuristic policy and that of the optimal policy. As common in the revenue management literature, the effectiveness of heuristic polices are evaluated in an asymptotic regime where resource capacities and customer arrivals are both scaled proportionally by a factor of k $(k = 1, 2, \ldots)$. Intuitively, this asymptotic regime increases market size while keeping resource scarcity, i.e., the ratio of capacity to demand, at a constant level. We assume this standard asymptotic setting throughout the chapter.

One popular heuristic for the NRM problem that is extensively studied in the academic literature and widely used in practice is based on the deterministic linear programming (DLP) approximation, where the customer demand distributions are replaced by their expectations. The solution of the DLP can then be used to construct heuristic policies. Under the asymptotic scaling defined above, Gallego and van Ryzin (1994) and Gallego and Ryzin (1997) have shown that the revenue loss of DLP-based static control policies is $\Theta(\sqrt{k})$ when the system size is scaled by k. The book by Talluri and Van Ryzin (2004) provides a comprehensive overview of different types of DLP-based control policies such as booking limit control, bid-price control, etc., and their variations.

An apparent weakness of the DLP approximation is that it ignores randomness in the arrival process and fails to incorporate information acquired through time. To include updated information, a simple approach is to re-optimize the DLP from time to time, while replacing the initial capacity in the DLP with the remaining capacity at each re-solving point. The new solution to the updated DLP is then used to adjust control policies. The re-solving approach is intuitive and widely used in practice. We refer to this family of solution techniques as *re-solving heuristics*. One might expect that re-solving the DLP would yield better performance since it includes updated information. Surprisingly, Cooper (2002) provides a counter-example where the performance of booking limit control deteriorates by re-solving the DLP. Furthermore, Chen and Homem-de-Mello (2010) give an example where re-solving the DLP worsens the performance for bid-price control. Jasin and Kumar

(2013) analyze the performance of re-solving of both booking limit and bid-price controls. They showed that when the initial capacity and customer arrival rates are both scaled by k, the revenue loss of re-solving heuristics is $\Omega(\sqrt{k})$, even by optimizing over the re-solving schedule or increasing re-solving frequency.

Despite those negative results, we note that there are several ways to construct control policies from the DLP, so it is possible that some control policies are suitable for applying the re-solving technique, while others are not. Some recent literature draws attention to a specific type of control policy called *probabilistic allocation*, which seems suitable for applying the re-solving technique. Probabilistic allocation control is a randomized algorithm that accepts each arriving customer with some probability. Using the probabilistic allocation control, Reiman and Wang (2008) propose a heuristic policy that re-solves the DLP exactly once during the horizon. In their proposed policy, the re-solving time is random and determined endogenously by the heuristic policy. In the asymptotic setting, Reiman and Wang (2008) show that the revenue loss of their policy is $o(\sqrt{k})$. This is an improvement over the $\Theta(\sqrt{k})$ revenue loss of DLP-based static policies.

Jasin and Kumar (2012) consider the NRM problem with customer choice, which generalizes the quantity-based NRM problem. They analyze an algorithm that is based on probabilistic allocation control and re-solves the DLP after each unit of time. They show the algorithm has a revenue loss of $O(1)$ when the system size is scaled by $k \to \infty$. A similar $O(1)$ revenue loss is obtained by Wu, Srikant, Liu, and Jiang (2015) for the case of one resource. However, both Jasin and Kumar (2012) and Wu, Srikant, Liu, and Jiang (2015)'s results require the optimal solution to DLP (before any updating) to be nondegenerate; this assumption will be formally stated in Section 2.3, which seems to be central to the hardness of the NRM problem. Moreover, Wu, Srikant, Liu, and Jiang (2015) show that when the optimal solution is nondegenerate but nearly degenerate, the constant factor in $O(1)$ can become arbitrarily large. In this chapter, we aim to establish a uniform $O(1)$ loss for the general NRM problem without assuming nondegeneracy.

2.1.2 Main contributions

We propose a new re-solving heuristic that has a uniformly bounded revenue loss when the system size is scaled by $k \to \infty$. (Recall that the rate of revenue loss is defined for a sequence of problems indexed by $k = 1, 2, \ldots$, where the capacities and arrival rates are multiplied by k, while other parameters are treated as constants.) The bound is uniform in the sense that it does not depend on the ratio between capacities and time. Therefore, this result does not require the nondegeneracy assumption. Our $O(1)$ bound improves the $o(\sqrt{k})$ bound in Reiman and Wang (2008), and also improves the $O(1)$ bound in Jasin and Kumar (2012), where the constant factor requires nondegeneracy assumption and depends implicitly on problem instances. (However, as we noted before, Jasin and Kumar (2012) considered the NRM problem with customer choice, which generalizes the quantity-based NRM problem.) We call our new algorithm Infrequent Re-solving with Thresholding (IRT). The intuition behind the IRT algorithm is that it is not necessary to update the DLP at early stage of the horizon, as the solution to the DLP barely changes after updating. It is sufficient to re-solve the DLP at a few carefully selected time points near the *end* of the horizon. In total, the IRT algorithm has $O(\log\log k)$ re-solving times for a system with scaling size k. Furthermore, a "thresholding" technique is applied in case that the DLP solution after re-solving is nearly degenerate. The re-solving schedule and the thresholds of the IRT algorithm are designed in such a way that the accumulated random deviations before the re-solving point can be corrected after re-solving with high probability.

Then, we give a tight performance bound of the re-solving heuristic proposed by Jasin and Kumar, 2012, but without assuming the optimal solution to the DLP is nondegenerate. The heuristic in Jasin and Kumar, 2012, which we call Frequent Re-solving (FR), re-solves the DLP after each unit of time. One would expect that by re-solving the DLP frequently and thus constantly updating capacity information, the decision maker can improve the expected revenue. Indeed, Jasin and Kumar (2012) have shown that under the nondegeneracy assumption, the revenue loss of this policy is $O(1)$ when the system size is scaled by

$k \to \infty$. However, we find that the revenue loss of this policy is $\Theta(\sqrt{k})$ in general, which has the same order of revenue loss as DLP-based static heuristics without any re-solving (Gallego and Ryzin, 1997; Talluri and Ryzin, 1998; Cooper, 2002). In particular, Proposition 2 shows that there exists a problem instance where the revenue loss of this policy is at least $\Omega(\sqrt{k})$. To analyze this instance, we used the Berry-Esseen bound and Freedman's inequality to show that the probability of revenue loss being larger than $\Omega(\sqrt{k})$ is bounded away from 0. This result suggests that the nondegeneracy assumption made by Jasin and Kumar (2012) is necessary to obtain $O(1)$ revenue loss, and explains why the $O(1)$ factor in Wu, Srikant, Liu, and Jiang, 2015 must be arbitrarily large when the DLP optimal solution is converging to a degenerate point. Then, Proposition 3 shows that the revenue loss of this policy is bounded above by $O(\sqrt{k})$ in the general case, which also improves the $o(k)$ bound in Maglaras and Meissner, 2006. The proof is based on a key inequality that bounds the average remaining capacity as a function of the remaining time.

In Figure 2.2, we summarize the performance of existing re-solving heuristics for the NRM problem. In this figure, the vertical axis represents the expected revenue which increases from the bottom to the top. We highlight the gap between different heuristics and upper bounds compared to the optimal revenue, which in principle can be obtained from dynamic programming but is hard to compute directly. The main result of this chapter (Theorem 1) simultaneously establishes an $O(1)$ upper bound of the hindsight optimum and an $O(1)$ revenue loss of the IRT algorithm.

2.1.3 Other related work

The re-solving heuristics defined in the NRM context is generally known as *certainty equivalent control* in dynamic programming. In certainty equivalent control, each random disturbance is fixed at a nominal value (e.g., its mean), and then an optimal control sequence for the certainty equivalence approximation is found. Only the first control in the sequence is applied, the rest of them are discarded, and the same process is repeated in the next

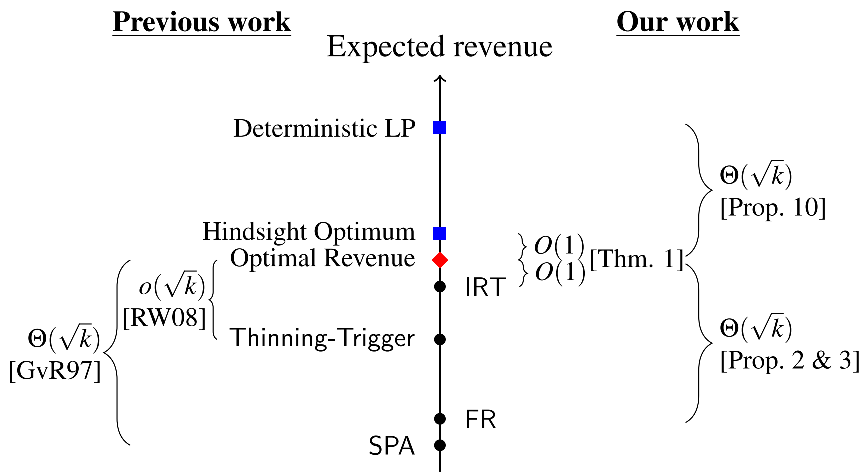


Figure 2.2: Summary of the results in the previous literature (on the left side) and our main results (on the right side). The red node (●) represents the expected revenue of the optimal policy (hard to compute); the blue nodes (●) represent upper bounds to the optimal revenue; and the black nodes (●) refer to revenues earned under different heuristics. The factor k is the scale of both time horizon and capacities.

stage. An introduction to certainty equivalent control can be found in Bertsekas (2005, Section 6.1). Secomandi (2008) discussed whether certainty equivalent control guarantees performance improvement in the network revenue management setting.

The quantity-based NRM model can be generalized in several ways. One extension assumes that the decision maker offers a set of products to each arriving customer, and customers choose some products from the offered set based on some discrete choice model (Talluri and van Ryzin, 2004; Liu and Ryzin, 2008). Another stream of literature assumes that either the customers' arrival process or the distribution of their reservation price is unknown, and requires the decision maker to learn the distribution exclusively from past observations (Besbes and Zeevi, 2012; Jasin, 2015; Ferreira, Simchi-Levi, and Wang, 2018). Talluri and Van Ryzin, 2004; Maglaras and Meissner, 2006 discussed the case where the decision maker posts price (price-based NRM) versus the case where the decision maker chooses accept/reject (quantity-based NRM).

The NRM problem considered here is related to the online knapsack/secretary problem studied by Kleywegt and Papastavrou (1998), Kleinberg (2005), and Babaioff, Immorlica, Kempe, and Kleinberg (2007), Arlotto and Gurvich (2019), and Arlotto and Xie (2020). In particular, Arlotto and Gurvich (2019) considers a multi-selection secretary problem, where the decision maker sequentially selects i.i.d. random variables in order to maximize the expected value of the sum given a fixed budget. As such, by viewing each random variable as a customer arrival, the multi-selection secretary problem is a special case of the NRM problem in which there is only a single resource and each customer requests exactly one unit of the resource. Arlotto and Gurvich (2019) proposes an online policy that has a uniformly bounded regret compared to the optimal offline policy. Their policy accepts or rejects an arriving customer by comparing the budget ratio, i.e., ratio of remaining budget to remaining arrivals, to some fixed thresholds. However, it is unclear whether their technique can be generalized to the general NRM setting with multiple resources, since the thresholds in their policy are specifically defined for a single resource.

Recently, Vera and Banerjee (2021) studies an online packing problem, which has the same mathematical formulation as the network revenue management problem. They propose a re-solving heuristic that achieves $O(1)$ revenue loss without the nondegeneracy assumption and under mild assumptions on the customer arrival processes. Unlike the IRT algorithm, their proposed algorithm re-solves the DLP every time there is an arrival; the algorithm then accepts that arrival if the acceptance probability from the DLP is greater than 0.5 and rejects it otherwise. Their proof is based on a novel argument that compensates the optimal offline algorithm and forces it to follow the decisions of their online algorithm. The design of their algorithm and their proof idea are significantly different from those in this paper.

2.1.4 Notation

For a positive integer n, let $[n]$ denote the set $\{1,\dots,n\}$. Given two real numbers $a \in \mathbb{R}$ and $b \in \mathbb{R}$, let $a \wedge b := \min\{a,b\}$, $a \vee b := \max\{a,b\}$, and $a^+ := a \vee 0$. For any real number x, let $\lfloor x \rfloor$ be the largest integer less than or equal to x, and let $\lceil x \rceil$ be the smallest integer greater than or equal to x. For a set S, let $|S|$ denote the cardinality of S. For two functions $f(T)$ and $g(T) > 0$, we write $f(T) = O(g(T))$ if there exists a constant M_1 and a constant T_1 such that $f(T) \leq M_1 g(T)$ for all $T \geq T_1$; we write $f(T) = \Omega(g(T))$ if there exists a constant M_2 and a constant T_2 such that $f(T) \geq M_2 g(T)$ for all $T \geq T_2$. If $f(T) = O(g(T))$ and $f(T) = \Omega(g(T))$ both hold, we denote it by $f(T) = \Theta(g(T))$.

2.2 Problem Formulation and Approximations

Suppose there is a finite horizon with length T. There are n classes of customers indexed by $j \in [n]$. The arrival process of customers in class j, $\{\Lambda_j(t), 0 \leq t \leq T\}$, follows a Poisson process of rate λ_j. We let $\Lambda_j(t_1,t_2)$ denote the number the arrivals of class j customers during $(t_1,t_2]$ for $0 \leq t_1 < t_2 \leq T$, i.e., $\Lambda_j(t_1,t_2) = \Lambda_j(t_2) - \Lambda_j(t_1)$. Arrival processes of different classes are independent. Upon arrival, each customer must either be accepted or rejected. Let r_j denote the revenue received by accepting a class j customer and $r = [r_1,\ldots,r_n]^\top$ be the vector of such revenues. There are m resources indexed by $l \in [m]$, where resource l has initial capacity C_l. The vector of the initial capacities is given by $C = [C_1,\ldots,C_m]^\top$. If a customer is accepted, a_{lj} units of resource l is consumed to serve a class j customer; let $A_j = [a_{1j},\ldots,a_{mj}]^\top$ be the column vector associated with class j customers. Let $A \in \mathbb{R}^{m\times n}$ be the *bill-of-materials* (BOM) matrix defined as $A = [A_1;\ldots;A_n]$. If a customer is rejected, no revenue is collected and no resource is used. Unused resources at the end of the horizon are perishable and have no salvage value. The objective of the decision maker is to maximize the expected revenue earned during the entire horizon by deciding whether or not to accept each arriving customer.

For a control policy π, let $z_j^\pi(t_1,t_2)$ be the number of class j customers admitted during $(t_1,t_2]$ ($\forall j \in [n], 0 \leq t_1 < t_2 \leq T$) under that policy. We call a policy *admissible* if it is non-anticipating and satisfies

$$\sum_{j=1}^{n} A_j z_j^\pi(0,T) \leq C \text{ a.s.,} \quad \text{and} \quad z_j^\pi(t_1,t_2) \leq \Lambda_j(t_1,t_2) \text{ a.s., } \forall j \in [n], 0 \leq t_1 < t_2 \leq T.$$

Let Π be the set of all admissible policies. The expected revenue under policy $\pi \in \Pi$ is defined as $v^\pi = \mathbb{E}\left[\sum_{j=1}^n r_j z_j^\pi(0,T)\right]$. We use $v^* = \sup_{\pi\in\Pi} v^\pi$ to denote the expected revenue under the optimal policy. If v^π is the expected revenue of a feasible policy $\pi \in \Pi$, we call $v^* - v^\pi$ the *revenue loss* of policy π.

2.2.1 Asymptotic framework

The standard asymptotic framework in revenue management measures performance of heuristics when the capacities and customer arrivals are scaled up proportionally. Under this asymptotic scaling, we consider revenue loss of a sequence of problems, indexed by $k = 1,2,\ldots$, where the capacities and arrival rates are multiplied by k, while all other problem parameters are treated as constants.

To avoid cumbersome notation where lots of variables and quantities are indexed by k, in the rest of the paper, we consider a different but equivalent asymptotic scaling, where the customer arrival rates λ_j ($j \in [n]$) are kept as constants, the time horizon is scaled up by $T = 1,2,\ldots$, and the resource capacities are scaled up proportionally by $C_l = b_l T$ ($l \in [m]$). Since the arrivals follow Poisson processes, scaling up the arrival rates and scaling up the horizon length have the same effect. We will thus express the revenue loss of heuristics in the order of T. Note that the horizon length (T) plays the same role as the scaling factor (k) in the standard asymptotic regime. For example, if we say the revenue loss of an algorithm is $O(\sqrt{T})$, it implies that revenue loss of that algorithm is $O(\sqrt{k})$ under the standard scaling regime.

2.2.2 Previous work on upper bound approximations

Deterministic linear program (DLP).

The DLP formulation is obtained by replacing all random variables with their expectations. As the expected number of arrivals of class j customers during the horizon is $\lambda_j T$ for

$j \in [n]$, the DLP formulation is given by

$$v^{\mathrm{DLP}} = \max_{y} \Big\{ \sum_{j=1}^{n} r_j y_j \Bigm| \sum_{j=1}^{n} A_j y_j \leq C, \text{and } 0 \leq y_j \leq \lambda_j T,\ \forall j \in [n] \Big\}. \tag{2.1}$$

In this formulation, decision variables y_j can be viewed as the expected number of class j customers to be accepted in $[0,T]$. The first constraint specifies that the expected usage of all m resources cannot exceed their initial capacities, $C = [C_1, \ldots, C_m]^\top$, and the second constraint specifies that the number of accepted customers from class j cannot exceed the expected number of arrivals, $\lambda_j T$.

Suppose y^* is an optimal solution to (2.1). The optimal value of DLP is given by $v^{\mathrm{DLP}} = \sum_{j=1}^{n} r_j y_j^*$. It can be shown that v^{DLP} is an upper bound of the expected revenue of the optimal policy, v^*, namely $v^* \leq v^{\mathrm{DLP}}$ (Gallego and Ryzin, 1997). Intuitively, DLP is a relaxation of the original problem since it only requires the capacity constraints to be satisfied in expectation, so v^{DLP} is an upper bound of v^*.

Equivalently, we can reformulate the DLP in (2.1) by letting x_j be the average number of class j customers accepted per unit time, i.e., $x_j = y_j/T$. Then, we get

$$v^{\mathrm{DLP}} = \max_{x} \Big\{ T \sum_{j=1}^{n} r_j x_j \Bigm| \sum_{j=1}^{n} A_j x_j \leq b, \text{ and } 0 \leq x_j \leq \lambda_j,\ \forall j \in [n] \Big\}, \tag{2.2}$$

where $b = [b_1, \ldots, b_m]^\top$ refers to the vector of available resources per unit time, i.e., $b_l = C_l/T, \forall l \in [m]$. Let x_j^* for $j \in [n]$ be an optimal solution to (2.2). The optimal value to the DLP is given by $v^{\mathrm{DLP}} = T \sum_{j=1}^{n} r_j x_j^*$.

Hindsight optimum.

The hindsight optimum is the optimal revenue obtained when the total number of arrivals is known in advance. Recall that the random variable $\Lambda_j(T)$ represents the total arrivals of class j customers in $[0,T]$. If the values of $\Lambda_j(T)$ are known, let z_j be the number of class

j customers accepted in $[0,T]$; the optimal acceptance policy is given by

$$V^{\mathrm{HO}} = \max_{z} \left\{ \sum_{j=1}^{n} r_j z_j \;\middle|\; \sum_{j=1}^{n} A_j z_j \le C, \text{ and } 0 \le z_j \le \Lambda_j(T),\ \forall j \in [n] \right\}. \tag{2.3}$$

Let V^{HO} be the optimal objective value and $\bar{z}_j,\ j \in [n]$ be the optimal solution; note that V^{HO} and $\bar{z}_j$'s are random variables that depend on $\Lambda_j(T)$. The *hindsight optimum* (HO) is defined as the expectation of the optimal objective value, i.e, $v^{\mathrm{HO}} = \mathbb{E}[V^{\mathrm{HO}}] = \mathbb{E}[\sum_{j=1}^{n} r_j \bar{z}_j]$.

The hindsight optimum is obviously an upper bound to the optimal revenue of the original problem, since the decision maker does not know the future arrivals at time $t = 0$. In fact, it can be shown that hindsight optimum is a *tighter* upper bound than the DLP, namely $v^* \le v^{\mathrm{HO}} \le v^{\mathrm{DLP}}$ (Talluri and Ryzin, 1998). This is easily verified since the expectation of the hindsight optimal solution, $\mathbb{E}[\bar{z}_j]$, is a feasible solution to the DLP.

We use the following definition throughout the paper.

Definition 1. *Let v^{π} be the expected revenue associated with an admissible control policy π. We refer to $v^{\mathrm{HO}} - v^{\pi}$ as the* regret *of that policy. (Note: since $v^* \le v^{\mathrm{HO}}$, the revenue loss of the control policy, $v^* - v^{\pi}$, is upper bounded by its regret.)*

2.2.3 Static probabilistic allocation heuristic

There are various ways to construct heuristic policies using the optimal solution of DLP. An overview can be found in Talluri and Van Ryzin (2004, Ch. 2). One intuitive approach is to interpret the solution to DLP as acceptance probabilities. Suppose x^* is an optimal solution to DLP in (2.2). For each arriving customer, if the customer belongs to class j, s/he would be accepted independently with probability x_j^*/λ_j throughout the time horizon. Since customers from each class are accepted with probabilities that are static, we call this heuristic Static Probabilistic Allocation (SPA). The SPA policy is formally stated in Algorithm 1.

The expected revenue of the SPA policy, denoted by v^{SPA}, can be computed as follows.

Algorithm 1 Static probabilistic allocation heuristic: SPA

initialize $x^* \leftarrow \arg\max_x \left\{ \sum_{j=1}^n r_j x_j \;\middle|\; \sum_{j=1}^n A_j x_j \leq C/T, \text{ and } 0 \leq x_j \leq \lambda_j, \forall j \in [n] \right\}$;
$C' \leftarrow C$
for all customers arriving in $[0,T]$ **do**
  if the customer belongs to class j and $A_j \leq C'$ ($\forall j \in [n]$) **then**
    accept the customer with probability x_j^*/λ_j
    if the customer is accepted, update capacity $C' \leftarrow C' - A_j$
  else
    reject the customer

Since the total number of arrivals from class j follows a Poisson distribution with mean $\lambda_j T$, the number of customers that the algorithm *attempts* to accept from class j follows a Poisson distribution with mean $(\lambda_j T) \cdot x_j^*/\lambda_j = x_j^* T = y_j^*$. Due to limited capacity, we must reject any customer from class j if the remaining capacity C' does not satisfy $A_j \leq C'$. It is straightforward to show that the expected number of customers who are turned away due to capacity limits is $O(\sqrt{T})$ (see e.g. Gallego and Ryzin, 1997; Reiman and Wang, 2008). Thus, we have

$$v^{\mathsf{SPA}} = \sum_{j=1}^n r_j y_j^* - O(\sqrt{T}) = v^{\mathrm{DLP}} - O(\sqrt{T}).$$

Recall from §2.2.2 that v^{DLP} is an upper bound of the expected revenue under the optimal policy, namely $v^* \leq v^{\mathrm{DLP}}$. Thus, the revenue of SPA is bounded by $v^{\mathsf{SPA}} \geq v^* - O(\sqrt{T})$.

2.3 Frequent Re-solving and Degeneracy

An obvious drawback of the SPA policy constructed from the DLP is that it does not take into account the randomness of demand or the updated information after $t = 0$. This motivates us to consider *re-solving heuristics*, which periodically re-optimize the DLP using the updated capacity information to adjust customer admission controls.

In particular, the following re-solving heuristic, which we referred to as Frequent Re-solving (FR), has been studied by Jasin and Kumar (2012) and Wu, Srikant, Liu, and Jiang (2015). The FR policy divides the horizon into T periods and re-solves the LP at the beginning of each period. At time $t = 0, 1, \ldots, T-1$, let $C_l(t)$ denote the remaining capacity of resource $l \in [m]$. We let $b_l(t) := \frac{C_l(t)}{T-t}$ be the average available capacity of resource l in period t. Let $C(t)$ and $b(t)$ denote the vectors of the remaining capacities and the average remaining capacities per unit time at time t, respectively, for all the resources. We outline the FR policy in Algorithm 2.

Algorithm 2 Frequent Re-solving Heuristic: FR

initialize: set $C(0) = C$ and $b(0) = C/T$
for $t = 0, 1, \ldots, T-1$ **do**
  set $x(t) \leftarrow \arg\max_x \left\{ \sum_{j=1}^n r_j x_j \;\middle|\; \sum_{j=1}^n A_j x_j \leq b(t), \text{ and } 0 \leq x_j \leq \lambda_j, \forall j \in [n] \right\}$
  set $C' \leftarrow C(t)$
  for all customers arriving in $[t, t+1)$ **do**
    if the customer belongs to class j and $A_j \leq C'$ $(\forall j \in [n])$ **then**
      accept the customer with probability $x_j(t)/\lambda_j$
      if the customer is accepted, update $C' \leftarrow C' - A_j$
    else
      reject the customer
  set $C(t+1) \leftarrow C'$ and $b(t+1) \leftarrow \frac{C(t+1)}{T-t-1}$

Jasin and Kumar (2012) show that when the optimal solution to DLP (2.2) is *nondegenerate*, FR has a revenue loss of $O(1)$, namely, the revenue loss is bounded when the problem size k grows. The optimal solution x^* is nondegenerate if

$$\left|\{j \in [n] : x_j^* = 0 \text{ or } x_j^* = \lambda_j\}\right| + \left|\{l \in [m] : \sum_{j=1}^n a_{lj} x_j^* = b_j\}\right| = n. \tag{2.4}$$

The $O(1)$ loss is a significant improvement from the $O(\sqrt{T})$ revenue loss of SPA.

However, the assumption of nondegenerate DLP solution is critical to achieve the $O(1)$ loss. The proofs by Jasin and Kumar, 2012 and Wu, Srikant, Liu, and Jiang, 2015 are built on a key observation that the ratio of remaining capacities to remaining time, $b(t)$, is a martingale (see also Arlotto and Gurvich (2019) for a discussion on this martingale property). If the optimal solution x^* is safely far from any degenerate solutions, with high probability, the adjusted solution $x(t)$ in Algorithm 2 shares the same basis with x^*, so the revenue loss of FR can be bounded. It is unclear from the analysis of Jasin and Kumar

(2012) and Wu, Srikant, Liu, and Jiang (2015) whether the nondegeneracy assumption is just an artifact of their analysis technique or something intrinsic to the performance of FR. This motivates us to examine closely the role of the nondegeneracy assumption.

2.3.1 A degenerate example

We will illustrate the issue of degenerate DLP solutions using the following numerical example, while deferring the theoretical analysis of the FR policy to Section 2.5.

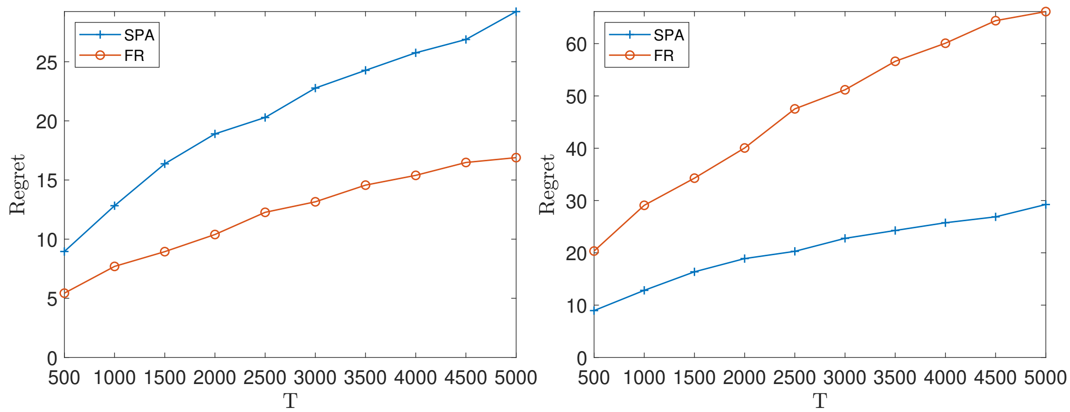
Suppose there are two classes of customers and one resource. Customers from each class arrive according to a Poisson process with rate 1. Customers from both classes, if accepted, consume one unit of resource, but pay different prices, r_1 and r_2. First, we compare the expected revenue loss of the FR policy and the SPA policy, which does not re-solve after $t = 0$, to examine the effect of frequent re-solving. We simulate the FR policy and the SPA policy when the average capacity per unit time $b = 1$ (so the total capacity is T) for two price scenarios: (a) $r_1 = 2$ and $r_2 = 1$; (b) $r_1 = 5$ and $r_2 = 1$ and for varying horizon length $T = 500, \ldots, 5000$. In both scenarios, the optimal solution to the DLP (2.2) is $x_1^* = 1, x_2^* = 0$. From Equation (2.4), we have

$$\left|\{j \in [n] : x_j^* = 0 \text{ or } x_j^* = \lambda_j\}\right| + \left|\{l \in [m] : \sum_{j=1}^{n} a_{lj} x_j^* = b_j\}\right| = 3 > n = 2,$$

thus the DLP solution in this example is degenerate.

Recall that the expected revenue loss of the FR policy is defined as $v^* - v^{\mathsf{FR}}$. Since calculating v^* requires solving dynamic programs, we use the regret $v^{\mathrm{HO}} - v^{\mathsf{FR}}$ (see the definition in §2.2.2) as a proxy of the expected revenue loss. In §2.4, we will show that $v^{\mathrm{HO}} - v^* = O(1)$, so this substitution does not affect the rate of revenue loss. Figure 2.3 plots the regrets under the FR policy and the SPA policy over 1000 sample paths.

We make the following observations from Figure 2.3. First, while the revenue loss of FR in scenario (a) is lower than that obtained from applying the SPA policy, the relationship



(a) $r_1 = 2$ and $r_2 = 1$ (b) $r_1 = 5$ and $r_2 = 1$

Figure 2.3: Regret under the FR policy (with re-solving) and the SPA policy (without re-solving).

is reversed in scenario (b). In other words, re-solving the DLP does not always lead to better performance. The intuition behind this result is that when the ratio r_1/r_2 is large, such as in scenario (b), rejecting a Class 2 customer to save the capacity for a potential future Class 1 customer is more profitable. The SPA policy accepts every customer from Class 1 and rejects all customers from Class 2, since the solution to the DLP (without re-solving) is $x_1^* = 1, x_2^* = 0$. This static policy is indeed optimal when $r_1/r_2 \to \infty$. In contrast, the FR policy constantly adjusts accepting probabilities, and starts to accept Class 2 customers when the actual arrival of Class 1 customers falls below its average. Second, we observe from Figure 2.3 that the revenue losses of both SPA and FR seem to have the same growth rate as horizon length T increases. (It is well-known that the revenue loss of SPA is of order $\Theta(\sqrt{T})$; see §2.2.3 and Proposition 11 in Online Appendix A.2.) This result is in contrast with Jasin and Kumar (2012), which show that when the solution to the DLP is nondegenerate, the expected revenue loss of FR is $O(1)$. However, we note that the nondegeneracy assumption made by Jasin and Kumar does not hold in this example, since the DLP has a unique solution that is degenerate.

Next, we simulate the FR policy when $r_1 = 2$, $r_2 = 1$ and $T = 50000$ for varying average capacity per unit time $b = 0.5, \dots, 2$. Note that when $b = 1$ and $b = 2$, the optimal solutions to the DLP (2.2) are $x_1^* = 1, x_2^* = 0$ and $x_1^* = 1, x_2^* = 1$, respectively, which are degenerate according to Equation (2.4). When $b \neq 1$ and $b \neq 2$, the solution to the DLP is nondegenerate. Therefore, by changing the value of b, we can evaluate the performance of FR with either degenerate or nondegenerate DLP solutions. Figure 2.4 shows the regret under the FR policy over 1000 sample paths.

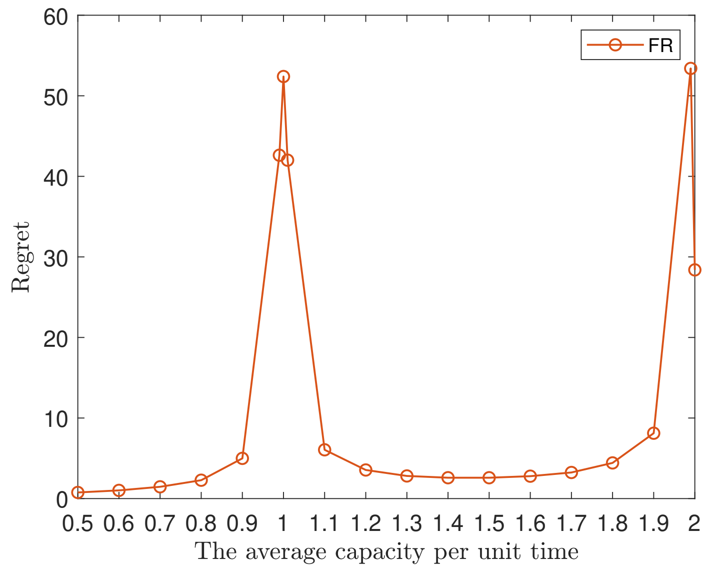
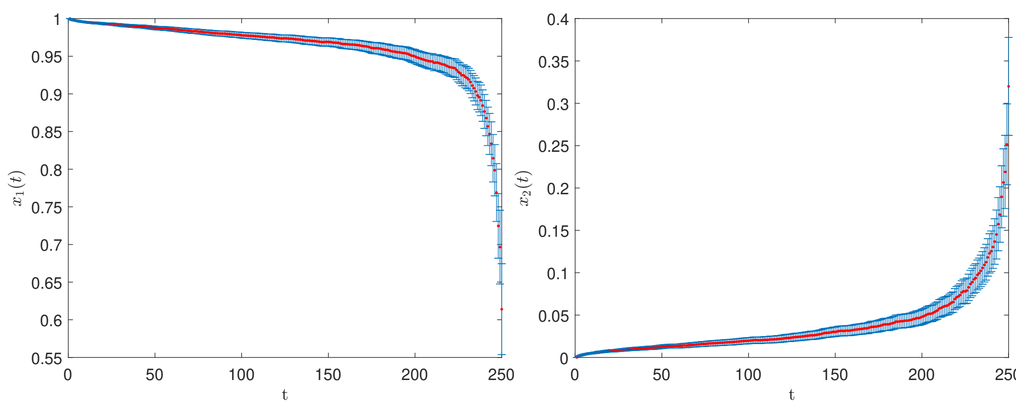


Figure 2.4: Regret under the FR policy for $r_1 = 2, r_2 = 1$ and $T = 50000$.

The simulation result from Figure 2.4 shows that the expected revenue loss under the FR policy is sensitive to the value of capacity rate b. When b is far away from the degenerate points (i.e., $b = 1$ and $b = 2$), FR performs well and has small revenue loss. However, the revenue loss increases significantly when the optimal DLP solution is close to degenerate (e.g., $b = 0.95$).

We notice that the observation from Figure 2.4 is consistent with the analysis by Jasin and Kumar, 2012. Even though Jasin and Kumar, 2012 proves that the revenue loss of FR is bounded by a constant whenever the DLP solution is nondegenerate, their analysis *does not imply the constant is uniform over all b's*. Rather, the constant bound from their analysis critically depends on the distance between b and its nearest degenerate point. When the optimal DLP solution is close to degenerate, the bound in Jasin and Kumar, 2012 can be arbitrarily large. Figure 2.4 shows that this phenomenon is not merely a consequence of the analysis technique from Jasin and Kumar, 2012, but reflects the actual performance of the FR policy.

Let us now turn our attention to the dynamics of the DLP solutions after updating under the FR policy. Figure 2.5 shows the trajectory of the DLP solutions of under FR when $r_1 = 2, r_2 = 1, b = 1$ and $T = 250$. In the figure, we simulated the policy 1000 times, and plotted the median, the 5th percentile, and the 95th percentile of the DLP solutions over the 1000 sample paths. It can be observed that the DLP solution barely changes near the beginning of the horizon, but changes significantly near the end of the horizon. These plots shed some light on the importance of each re-solving time: re-solving near the beginning of the horizon is not as important as re-solving near the end of the horizon.



(a) class 1 customer (b) class 2 customer

Figure 2.5: The 5th percentiles and 95th percentiles of the DLP solutions (i.e., acceptance probabilities) of the two customer classes under FR in each period for $T = 250$.

2.4 A Re-solving Heuristic with Uniformly Bounded Loss

In this section, we propose a new re-solving algorithm. The main result of this section is to show that this algorithm has uniformly bounded revenue loss given any horizon length T and starting capacity C, without requiring the nondegeneracy assumption.

2.4.1 Definition of the IRT algorithm

We propose an algorithm called Infrequent Re-solving and Thresholding (IRT). The IRT policy has two distinct features compared to the FR policy: 1) the DLP is *not* re-solved in every period; 2) customers acceptance probabilities are adjusted by some thresholds.

Unlike the FR policy, the IRT policy re-solves the DLP for only $O(\log\log T)$ times during a horizon of length T. The re-solving schedule is defined as follows. Given horizon length T, we set $K = \left\lceil \frac{\log\log T}{\log(6/5)} \right\rceil$. Let $\{t_u^*, \forall u \in [K]\}$ denote a sequence of re-solving times, where $\tau_u = T^{(5/6)^u}$ and $t_u^* = T - \tau_u$ for all $u \in [K]$. In addition, let $t_{K+1}^* = T$. Thus, the re-solving times divide the entire horizon into $K+1$ epochs: $[0, t_1^*)$, $[t_1^*, t_2^*)$, $\cdots$, $[t_K^*, t_{K+1}^*]$. Figure 2.6 illustrates the re-solving schedule of the IRT policy.

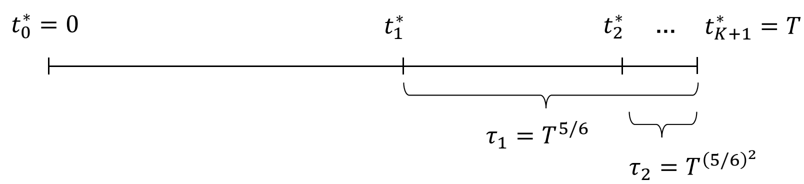


Figure 2.6: The re-solving times of the IRT policy is constructed recursively.

At the beginning of each epoch u $(0 \le u \le K)$, the algorithm solves an LP approximation to the dynamic programming problem—this LP is identical to the LP used in the FR algorithm, which uses information about remaining capacities and the mean of remaining customer arrivals. The optimal solution of the LP is then used to construct a probabilistic allocation control policy. The IRT policy applies thresholds to the allocation probabilities. In particular, in epoch $u \in \{0\} \cup [K-1]$ (except for the last epoch), the allocation probability for each class is rounded down to 0 if it is less than $\tau_u^{-1/4}$, or rounded up to 1 if it is larger than $1 - \tau_u^{-1/4}$. The complete definition of IRT is given in Algorithm 3.

Before we present the formal analysis of the IRT algorithm, it might be helpful to discuss the intuition behind the design of this algorithm. We start with the choice of the first re-solving time, t_1^*. The analysis by Reiman and Wang, 2008 shows that by setting $t_1^* \approx T - O(\sqrt{T})$, one re-solving of DLP is sufficient to reduce the regret to $O(T^{1/4})$. But

Algorithm 3 Infrequent Re-solving with Thresholding (IRT)

initialize: set $\tau_u = T^{(5/6)^u}$ and $t_u^* = T - \tau_u$ for all $u \in \{0\} \cup [K]$, where $K = \left\lceil \frac{\log\log T}{\log(6/5)} \right\rceil$

for $u = 0, 1, \ldots, K$ **do**

  set $x^u \leftarrow \arg\max_x \left\{ \sum_{j=1}^n r_j x_j \;\middle|\; \sum_{j=1}^n A_j x_j \le C(t_k^*)/\tau_k, \text{ and } 0 \le x_j \le \lambda_j, \forall j \in [n] \right\}$

  if $u < K$ **then**

    for $j \in [n]$ **do**

      if $x_j^u < \lambda_j \tau_u^{-1/4}$ **then**

        set $p_j^u \leftarrow 0$

      else if $x_j^u > \lambda_j (1 - \tau_u^{-1/4})$ **then**

        set $p_j^u \leftarrow 1$

      else

        set $p_j^u \leftarrow x_j^u / \lambda_j$

  else

    set $p_j^u \leftarrow x_j^u / \lambda_j$ for all $j \in [n]$

  set $C' \leftarrow C(t_u^*)$

  for $t \in [t_u^*, t_{u+1}^*)$ **do**

    observe requests from all arrival of customers

    if an arriving customer belongs to class j and $A_j \le C'$ $(\forall j \in [n])$ **then**

      accept the customer with probability p_j^u

      if accepted, update $C' \leftarrow C' - A_j$

    else

      reject the customer

  set $C(t_{u+1}^*) \leftarrow C'$

we note that if t_1^* is defined as in Reiman and Wang, 2008, additional re-optimizations after t_1^* cannot improve the regret rate. In the IRT algorithm, we choose the first re-solving time to be $t_1^* = T - T^{5/6}$, which is earlier than the re-solving time in Reiman and Wang, 2008. If no further re-solving is used, this policy leads to a regret rate of $O(T^{5/12})$ (Proposition 1). Even though the $O(T^{5/12})$ rate is worse than the $O(T^{1/4})$ rate in Reiman and Wang, 2008, as we choose an earlier re-solving time, more time is left for making further adjustments. Once we establish the $O(T^{5/12})$ regret rate with the first re-solving, we then use induction to prove that subsequent re-optimizations of the DLP can further reduce the regret, eventually reducing it to a constant. By definition, τ_u, the length of epoch u satisfies the recursive relationship $\tau_{u+1} = \tau_u^{5/6}$, $\forall u \in [K]$. This enables us to apply the induction hypothesis to epochs $u \geq 1$.

The $\tau_u^{-1/4}$ thresholds in the algorithm are critical to bounding the regret. As we have seen from the numerical example in §2.3.1, large losses can occur when the DLP solution is nearly degenerate. If we use a nearly degenerate solution to construct probabilistic allocation controls, some customer classes would have acceptance probabilities that are either very close to 0 or very close 1. As a result, the mean number of accepted or rejected customers is dominated by its standard deviation, making the control policy ineffective. More specifically, if the acceptance probability of class j customers is $\varepsilon \to 0$, the coefficient of variation of the number of customer accepted in one unit time is $1/\sqrt{\lambda_j \varepsilon} \to +\infty$. Therefore, if the acceptance probability of a customer class is almost 0, we might as well reject all customers from that class in the current epoch, as long as there is sufficient time left to accept customers in the next epoch. Similarly, if the acceptance probability of a customer class is almost 1, we might as well accept all customers from that class in the current epoch. This is the intuition behind adding thresholds to the acceptance probabilities in the IRT policy.

2.4.2 Analysis of the IRT policy

We now formally analyze the revenue loss (regret) of the IRT policy. The main result of this section is the following.

Theorem 1. *The regret of IRT policy defined in Algorithm 3 is bounded by* $v^{\mathsf{HO}} - v^{\mathsf{IRT}} = O(1)$*. The constant factor depends on the customer arrival rate* λ_j *(*$\forall j \in [n]$*), the revenues per customer* r_j *(*$\forall j \in [n]$*), and the BOM matrix A; however, this constant is independent of the time horizon T and the capacity vector* $\mathbf{C}$*.*

Theorem 1 states that the regret of IRT policy is $O(1)$. Moreover, this constant is independent of time horizon and capacities, so the performance of IRT is uniformly bounded when the capacity ratio C/T varies. Because degenerate DLP solution occurs only for some specific capacity ratios, the result in Theorem 1 does not require the nondegeneracy assumption in Jasin and Kumar, 2012.

Since the hindsight optimum v^{HO} is an upper bound of the expected revenue of the optimal policy v^*, we immediately get a bound on its revenue loss: $v^* - v^{\mathsf{IRT}} \leq v^{\mathsf{HO}} - v^{\mathsf{IRT}} = O(1)$. Moreover, Theorem 1 implied that hindsight optimum is a tight upper bound, satisfying $v^{\mathsf{HO}} - v^* \leq v^{\mathsf{HO}} - v^{\mathsf{IRT}} = O(1)$.

The complete proof of Theorem 1 can be found in Appendix §A.1.1. We outline the main idea of the proof here. In the proof, we define a sequence of auxiliary re-solving policies with increasing re-solving frequency. Recall that $K = \left\lceil \frac{\log\log T}{\log(6/5)} \right\rceil$ is the number of re-optimizations made by the IRT algorithm. For any $u \in [K]$, we define a policy that follows the IRT heuristic exactly in $[0, t_u^*)$, but then applies static allocation control in $[t_u^*, T]$. We refer to such a policy as IRT^u. Notice that when $u = K$, IRT^u coincides with IRT. Similarly, we define HO^u as a policy that is exactly the same as IRT in $[0, t_u^*)$ but applies the hindsight optimal policy in $[t_u^*, T]$. Our proof of Theorem 1 depends on the following proposition, proved in Appendix §A.1.2.

Proposition 1. *Given horizon length T, suppose the first re-solving time is $t_1^* = T - T^{5/6}$, then*

1. *the regret of* HO^1 *is* $O(Te^{-\kappa T^{1/6}})$*;*
2. *the regret of* IRT^1 *is* $O(Te^{-\kappa T^{1/6}}) + O(T^{5/12})$*.*

Here, we define $\kappa = \frac{\lambda_{\min}}{24(\alpha|J_\lambda|+1)^2}$*, where* $J_\lambda := \{j : x_j^* = \lambda_j\}$ *(recall that* x^* *is the solution to DLP),* $\lambda_{\min} := \min_{j\in[n]} \lambda_j$*, and* α *is a positive constant that depends on the BOM matrix* A*.*

Notice that IRT^1 is a non-anticipating and admissible policy, and its regret of $O(T^{5/12})$ is an improvement over the $O(\sqrt{T})$ bound of SPA. The policy HO^1 is not non-anticipating since it requires access to future arrival information; thus it is not practical and its sole purpose is to bound the performance of IRT^1 in the proof.

We then use Proposition 1 to prove Theorem 1 by induction. We illustrate the induction step using IRT^2, a policy that re-solves at $t_1^* = T - T^{5/6}$ and again at $t_2^* = T - T^{(5/6)^2}$. The regret of IRT^2 can be written as

$$\mathbb{E}[V^{\mathsf{HO}} - V^{\mathsf{IRT}^2}] = \underbrace{\mathbb{E}[V^{\mathsf{HO}} - V^{\mathsf{HO}^1}]}_{(*)} + \underbrace{\mathbb{E}[V^{\mathsf{HO}^1} - V^{\mathsf{HO}^2}]}_{(**)} + \underbrace{\mathbb{E}[V^{\mathsf{HO}^2} - V^{\mathsf{IRT}^2}]}_{(***)}.$$

The term $(*)$ is bounded by $O(Te^{-\kappa T^{1/6}})$ according to Proposition 1. For the term $(**)$, the policies HO^1 and HO^2 are identical up to time t_1^*. So applying part (1) of Proposition 1 to the subproblem in $(t_1^*, T]$, we get $\mathbb{E}[V^{\mathsf{HO}^1} - V^{\mathsf{HO}^2}] = O(T^{5/6}e^{-\kappa (T^{5/6})^{1/6}}) = O(T^{5/6}e^{-\kappa T^{5/36}})$. For the last term $(***)$, using the well-known result that static probabilistic allocation has a squared root regret, we have $\mathbb{E}[V^{\mathsf{HO}^2} - V^{\mathsf{IRT}^2}] = \mathbb{E}[V^{\mathsf{HO}}(t_2^*, T) - V^{\mathsf{SPA}}(t_2^*, T)] = O(\sqrt{T - t_2^*}) = O(T^{(5/6)^2/2})$. Combining these three terms, we get

$$v^{\mathsf{HO}} - v^{\mathsf{IRT}^2} = O(Te^{-\kappa T^{1/6}}) + O(T^{5/6}e^{-\kappa T^{5/36}}) + O(T^{25/72}) = O(T^{25/72}).$$

By induction, we show that if the decision maker re-solves for $K \geq 1$ times, where the

u-th ($u = 1, \cdots, K$) re-solving time is $t_u^* = T - T^{(5/6)^u}$, the regret is given by

$$v^{\mathsf{HO}} - v^{\mathsf{IRT}^K} = \sum_{u=0}^{K-1} O\left((T^{(5/6)^u} \exp\left(-\kappa T^{(5/6)^u/6}\right)\right) + O(T^{(5/6)^K/2}).$$

When $K = \left\lceil \frac{\log\log T}{\log(6/5)} \right\rceil$, the right-hand side of the above equation is bounded by a constant. In additional, the policy IRT^K is the same as IRT, so we prove that the regret of IRT is $v^{\mathsf{HO}} - v^{\mathsf{IRT}} = O(1)$.

2.4.3 Revisiting the degenerate example in Section 2.3.1

In Section 2.3.1, we considered a numerical example with two classes and one resource. We simulated the FR policy when $r_1 = 2$, $r_2 = 1$ and $T = 50000$ for varying average capacity $b = 0.5, \ldots, 2$, and showed that FR has poor performance when the DLP solution is either degenerate (i.e., $b = 1$ or $b = 2$) or nearly degenerate. We now test the IRT policy using the same example and compare it to the FR policy. Figure 2.7 plots the average regret under FR and IRT over 1000 sample paths.

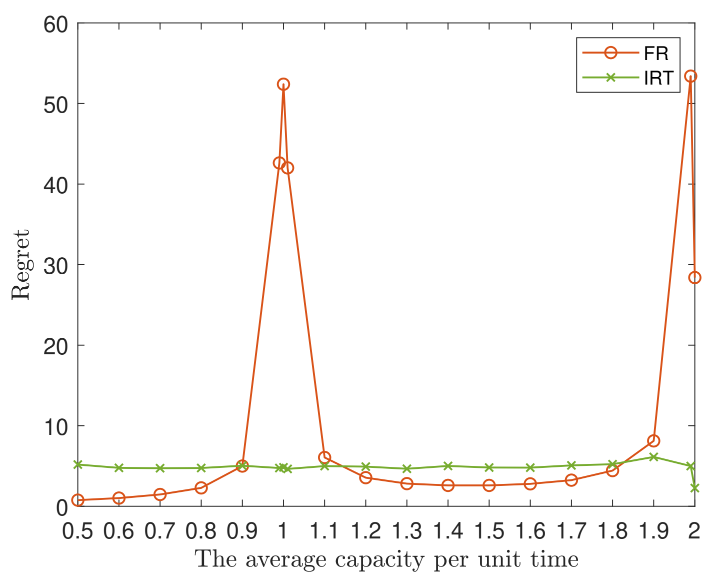


Figure 2.7: Regret under the FR policy and the IRT policy for $r_1 = 2, r_2 = 1$ and $T = 50000$.

It can be observed from Figure 2.7 that the regret under the proposed IRT policy is not sensitive to the average capacity per unit time. This result verifies Theorem 1 in that the regret of IRT is uniformly bounded with respect to the ratio between capacity and time. In contrast, the regret under the FR policy has two spikes that are associated with the two degenerate points ($b = 1$ and $b = 2$).

2.5 Analysis of the Frequent Resolving Policy

2.5.1 Lower bound of the revenue loss of FR

The simulation in Section 2.3.1 inspires us to analyze the performance of FR without the nondegeneracy assumption in order to gain a better understanding of the effect of frequent re-solving. First, we show that the regret under the FR policy is bounded below by $\Omega(\sqrt{T})$.

Proposition 2. *There exists a problem instance for which the regret of the FR policy defined in Algorithm 2 is bounded below by*

$$v^{\mathsf{HO}} - v^{\mathsf{FR}} = \Omega(\sqrt{T}).$$

Proposition 2 implies that the expected revenue loss under FR policy is bounded below by $\Omega(\sqrt{T})$ as well, because the revenue gap between the hindsight optimum (v^{HO}) and the optimal revenue (v^*) is $O(1)$ (Theorem 1). That is, we have

$$v^* - v^{\mathsf{FR}} = -(v^{\mathsf{HO}} - v^*) + v^{\mathsf{HO}} - v^{\mathsf{FR}} = -O(1) + \Omega(\sqrt{T}) = \Omega(\sqrt{T}).$$

To prove Proposition 2, we consider a problem instance with two classes of customers and one resource. We assume that customers from each class arrive according to a Poisson process with rate 1; the arrivals from two classes are independent. The initial resource capacity is T. Customers from both classes, if accepted, consume one unit of the resource, but pay different prices, $r_1 > r_2$. We consider the event when the number of class 1 customers that arrive during T period is more than T. If this event happens, the hindsight optimum will accept T of class 1 customers and none of class 2 customers. Conditional on that event, we use Freedman's inequality (Freedman (1975)) to show that with positive probability, the FR policy accepts $\Omega(\sqrt{T})$ of class 2 customers, and thus at most $T - \Omega(\sqrt{T})$ of class 1 customers. So the revenue of FR is at least $\Omega(\sqrt{T})$ less than the hindsight optimum. The complete proof can be found in Online Appendix §A.3.1.

2.5.2 Upper bound of the revenue loss of FR

In this section, we provide an upper bound of the expected revenue loss of the FR policy.

Proposition 3. *The gap between the expected revenue of the FR policy defined in Algorithm 2 and the optimal value of the* DLP *is bounded by*

$$v^{\text{DLP}} - v^{\text{FR}} = O(\sqrt{T}).$$

The constant pre-factor depends on the customer arrival rate λ_j ($\forall j \in [n]$), the revenues per customer r_j ($\forall j \in [n]$), and the BOM matrix A; however, it does not depend on the starting capacity C_l ($\forall l \in [m]$).

Since v^{DLP} is an upper bound of the expected revenue under the optimal policy (see Section 2.2.2), Proposition 3 immediately implies that the expected revenue loss of the FR policy when compared with the optimal revenue is bounded by $O(\sqrt{T})$. That is, $v^* - v^{\text{FR}} \leq v^{\text{DLP}} - v^{\text{FR}} = O(\sqrt{T})$. Combining Propositions 2 and 3 gives $v^* - v^{\text{FR}} = \Theta(\sqrt{T})$.

The proof of Proposition 3 can be found in Online Appendix §A.3.2. The proof is based on the following idea. Since the LP solved under the FR policy and the DLP (2.2) only differ in the right hand side of the capacity constraints, $b(t)$ and b, the expected revenue loss of the FR policy when compared to the optimal value of the DLP can be expressed in terms of $b(t)$ and b. More specifically, we show that the expected revenue loss during $[t, t+1)$ can be expressed as $O(\mathbb{E}[(b_l - b_l(t))^+])$ for each resource $l \in [m]$. Then, using the relationship between the average remaining capacity, $b(t)$, and the number of accepted customers up to time t, we prove that $O(\mathbb{E}[(b_l - b_l(t))^+]) = O(\frac{1}{\sqrt{T-t}})$. This completes the proof since $\sum_{t=0}^{T-1} O(\frac{1}{\sqrt{T-t}}) = O(\sqrt{T})$.

Although the $O(\sqrt{T})$ bound in Proposition 3 is looser than the $O(1)$ bound of FR in Jasin and Kumar (2012), it does not require the additional condition that the optimal solution to the DLP is nondegenerate. Given that the expected revenue loss of SPA is also $O(\sqrt{T})$ (see Online Appendix §A.2.2), we conclude that re-solving at least guarantees the same order of revenue loss compared to no re-solving.

2.6 Numerical Experiment

In this section, we evaluate the numerical performance of five different heuristics, which include

1. SPA: static probabilistic allocation heuristic (Algorithm 1)
2. FR: frequent re-solving heuristic (Algorithm 2)
3. IRT: infrequent re-solving with threhoslding (Algorithm 3)
4. IR: this algorithm uses the same re-solving schedule as IRT but without applying thresholding; i.e., the acceptance probability at iteration u is always set to $p_j^u \leftarrow x_j^u / \lambda_j$
5. FRT: this algorithm is motivated by IRT. We apply the same $\tau^{-1/4}$ thresholds from IRT to the frequent re-solving algorithm; see the complete description in Algorithm 4.

Recall that IRT has two distinct features compared to FR: it uses an infrequent re-solving schedule and adds thresholds for acceptance probabilities. The motivation to include IR and FRT in this test is to evaluate which of the two features plays a more important role.

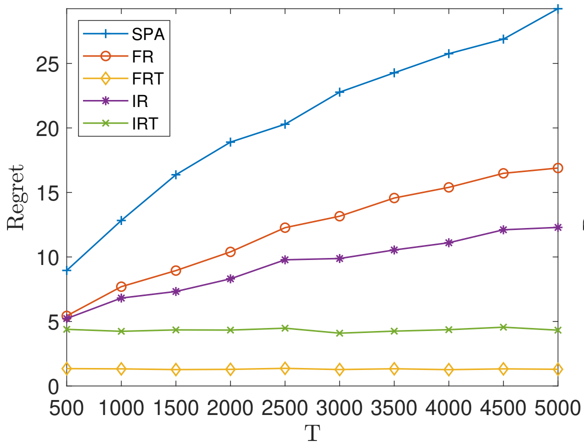
Algorithm 4 Frequent Re-solving with Thresholding: FRT

initialize: set $C(0) = C$ and $b(0) = C/T$
for $t = 0, 1, \ldots, T-1$ **do**
  set $x(t) \leftarrow \arg\max_x \left\{ \sum_{j=1}^n r_j x_j \;\middle|\; \sum_{j=1}^n A_j x_j \le b(t), \text{ and } 0 \le x_j \le \lambda_j, \forall j \in [n] \right\}$
  set $C' \leftarrow C(t)$
  for all customers arriving in $[t, t+1)$ **do**
    if the customer belongs to class j and $A_j \le C'$ ($\forall j \in [n]$) **then**
      if $x_j(t) < \lambda_j (T-t)^{-1/4}$ **then**
        reject the customer
      else if $x_j(t) > \lambda_j (1 - (T-t)^{-1/4})$ **then**
        accept the customer
      else
        accept the customer with probability $x_j(t)/\lambda_j$
      if the customer is accepted, update $C' \leftarrow C' - A_j$
    else
      reject the customer
  set $C(t+1) \leftarrow C'$ and $b(t+1) \leftarrow \frac{C(t+1)}{T-t-1}$

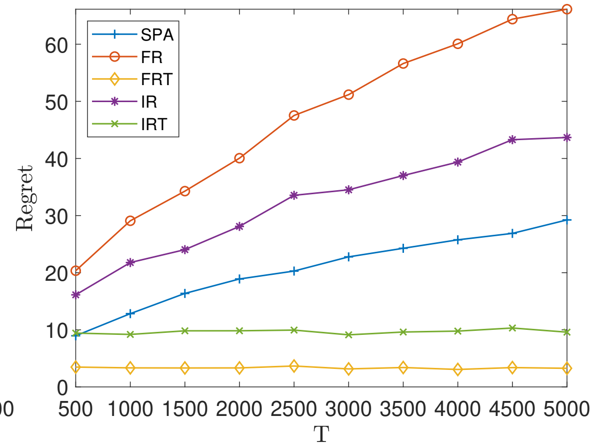
2.6.1 Single resource

We consider a revenue management problem with a single resource and two classes of customers. We assume that customers from each class arrive according to a Poisson process with rate 1. The arrivals of two classes are independent. Customers from both classes, if accepted, consume one unit of resource, but pay different prices, r_1 and r_2. We consider two cases: 1) $r_1 = 2, r_2 = 1$ and 2) $r_1 = 5, r_2 = 1$. We also test three settings for the average capacity per unit time: $b =$1, 1.1 and 1.5. When the average capacity per unit time is 1, the solution to the DLP is degenerate. The scenario where the average capacity is 1.1 represents a setting where the DLP solution is "nearly degenerate," and the scenario of 1.5 represents a setting where the DLP solution is far away from any degenerate point. We simulate the heuristics for two price and three average capacity per unit time scenarios defined above and for varying horizon length $T = 500, 1000, \ldots, 5000$.

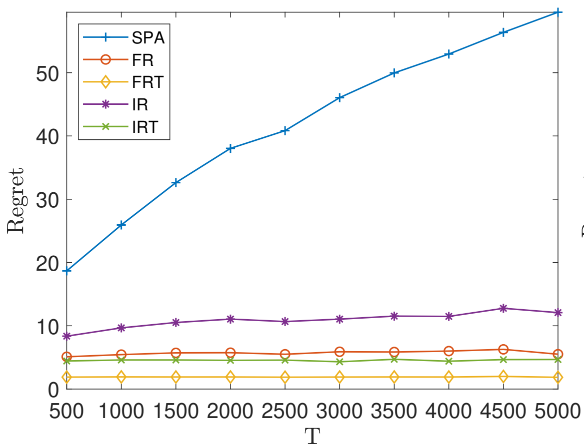
Figure 2.8 plots the regret under SPA, FR, FRT, IR, and IRT over 1000 sample paths. The first column shows the case when $r_1 = 2$ and $r_2 = 1$, while the second column shows the case when $r_1 = 5$ and $r_2 = 1$. The first, the second and the third rows illustrate the case when $b = 1$, $b = 1.1$, and $b = 1.5$ respectively. We make the following observations:



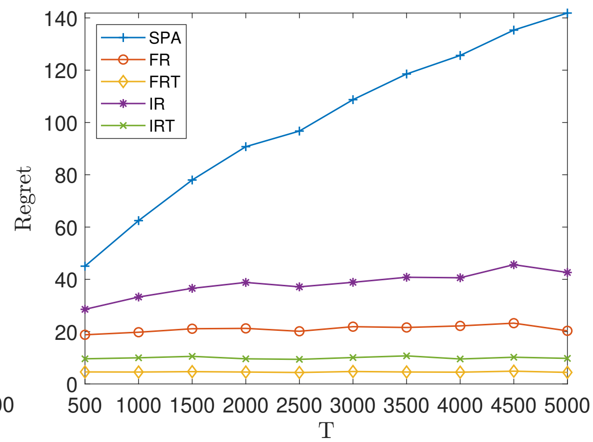
(a) $b = 1, r_1 = 2$ and $r_2 = 1$



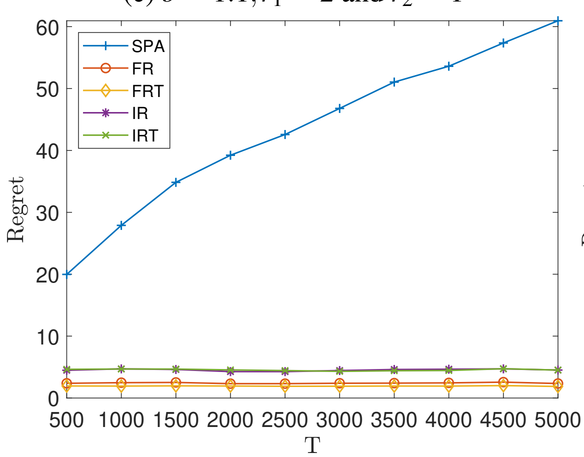
(b) $b = 1, r_1 = 5$ and $r_2 = 1$



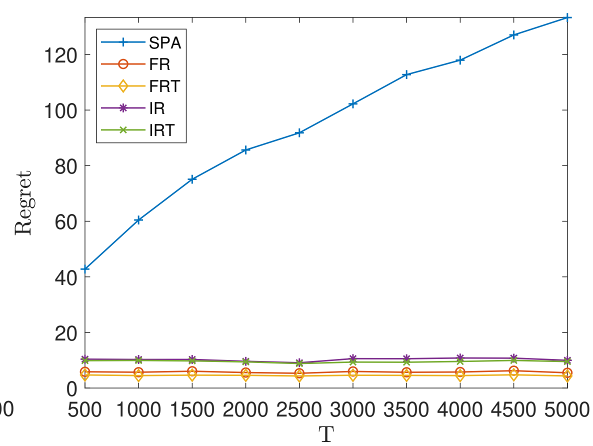
(c) $b = 1.1, r_1 = 2$ and $r_2 = 1$



(d) $b = 1.1, r_1 = 5$ and $r_2 = 1$



(e) $b = 1.5, r_1 = 2$ and $r_2 = 1$



(f) $b = 1.5, r_1 = 5$ and $r_2 = 1$

Figure 2.8: Regret under the SPA, the FR, the FRT, the IR, and the IRT policies for $T = 500, 1000, \ldots, 5000$.

1. When $r_1 = 2$ and $r_2 = 1$, the expected revenue loss under SPA is the largest for all average capacity per unit time and horizon length. This does not hold when $r_1 = 5$, $r_2 = 1$ and $b = 1$, where SPA is better than either frequent re-solving (FR) or infrequent re-solving (IR).

2. The expected revenue loss under IR is higher than the expected revenue loss under FR except when the problem is degenerate ($b = 1$). We conclude that choosing an infrequent re-solving schedule alone is not enough to achieve $O(1)$ loss.

3. The expected revenue losses under IRT and FRT remain constant for all cases as the horizon length increases. Moreover, although we don't have theoretical guarantee for FRT, the expected revenue loss under FRT often appears smaller than the expected revenue loss under IRT. This implies that appropriate thresholding is the main factor that leads to uniformly bounded regret for re-solving heuristics.

2.6.2 Multiple resources

Next, we consider a network revenue management problem with multiple resources. We consider the problem when there are five classes of customers and four types of resources. We assume that customers from each class arrive according to a Poisson process with rate 1; the arrivals of different classes are independent. The vector of the average capacities per unit time is given by $b = [1,1,1,1]^\top$. The vector of the revenue earned by accepting customers is given by $r = [10,3,6,1,2]^\top$. The bill-of-materials matrix is given by

$$A = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

We simulate the heuristics for varying horizon length $T = 500, 1000, \ldots, 5000$. Notice that in this example, the optimal solution to the DLP is degenerate.

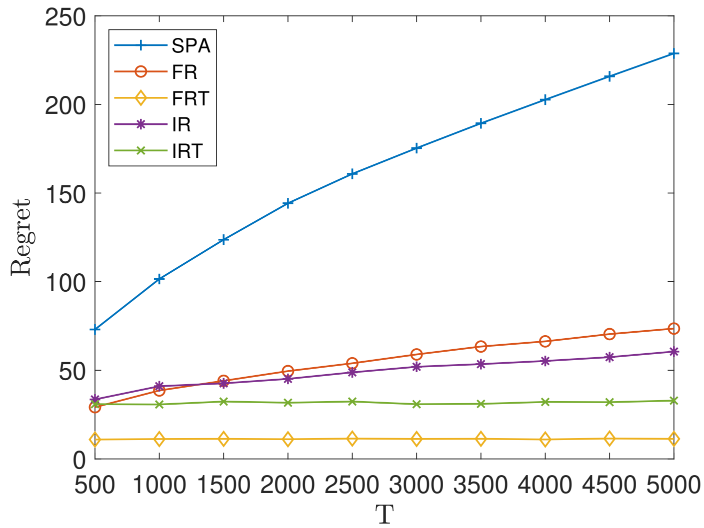


Figure 2.9: Regret under the SPA, the FR, the FRT, the IR, and the IRT when compared with the hindsight optimal for $T = 500, 1000, \ldots, 5000$.

Figure 2.9 plots the regrets under SPA, FR, FRT, IR, and IRT over 1000 sample paths. The result shows that the revenue losses of SPA scales poorly with horizon length T. In comparison, the revenue losses of FR and IR increase more slowly when T increases, and IR seems to perform slightly better for large T. The revenue losses of FRT and IRT remain constant as T grows. Moreover, the expected revenue loss under the IRT is higher than the expected revenue loss under the FRT. Again, this result implies that among the two factors, infrequent re-solving and thresholding, the latter plays a more important role.

2.7 Conclusion and Discussion

We study re-solving heuristics for the network revenue management (NRM) problem. A re-solving heuristic periodically re-optimizes a deterministic LP approximation of the original NRM problem. The main question considered in this paper is: can we find a simple and computationally efficient re-solving heuristic, whose expected revenue loss compared to the optimal policy is bounded by a constant even when both the time horizon and the resource capacities scale up?

We answer the above question in the affirmative by proposing a re-solving heuristic called Infrequent Re-solving with Thresholding (IRT), whose revenue loss is bounded by a constant independent of time horizon and resource capacities. This finding improves a previous result by Jasin and Kumar, 2012, showing that Frequent Re-solving (FR), an algorithm that re-solves the DLP after each unit of time, has $O(1)$ revenue loss, but requires the optimal solution to the DLP to be nondegenerate for the quantity-based NRM problem. Moreover, we show that when both time horizon and resource capacities scale up by $k = 1, 2, \ldots$, Frequent Re-solving (FR) has a revenue loss of $\Theta(\sqrt{k})$. This is a negative result, as most DLP-based heuristics can achieve the same revenue loss rate without using re-solving at all. However, we note that the analysis by Jasin and Kumar, 2012 considers customer choice behavior, while we focus on the classical quantity-based problem with independent demand model throughout this chapter. Jasin, 2014 and Atar and Reiman, 2012 also consider re-solving heuristics for the price-based NRM problem. It is unclear if one can extend our analysis technique to either the NRM problem with customer choice or the price-based NRM problem, because our proof is based on the notion of *hindsight optimum* (Reiman and Wang, 2008), which is not well-defined for the NRM problem with customer choice or the price-based NRM problem. We will leave this question for future research.

Our simulation results show that when the controls from FR are adjusted by some thresholds, the resulting algorithm FRT has very promising numerical performance and seems to have a bounded revenue loss as well. So far, we are not able to prove this result, mainly because the induction based proof we developed for IRT breaks down when the DLP is re-optimized every period. Recently, Vera and Banerjee, 2021 propose a different re-solving heuristic for the NRM problem, where the DLP is re-optimized every period and a fixed acceptance probability threshold of 0.5 is applied to every class for all periods. They show their heuristic also achieves $O(1)$ regret. Although the fixed threshold used by

Vera and Banerjee, 2021 is different from the time-varying thresholds we proposed in the FRT algorithm, we think their analysis technique may be helpful to establish the revenue loss bound of FRT.

CHAPTER 3

INTEGRATED PRICING AND ROUTING ON A NETWORK

3.1 Introduction

Most network flow models assume that demands on nodes are exogenous (Ahuja, Magnanti, and Orlin, 1993). In this chapter, we consider a generalization of the multicommodity network flow problem where demands are determined endogenously by setting prices on node pairs. This setup leads to an intricate interplay between pricing and network routing decisions: on the one hand, pricing decisions determine demand of flows and therefore affect routing decisions; on the other hand, routing decisions affect the cost of serving demand, which in turn informs how to optimally set the prices. In order to maximize overall system profit, the decision maker needs to consider the integration between pricing and routing decisions and optimize them jointly.

We consider a directed network represented by a set of nodes and arcs. Each arc has a fixed capacity and a linear cost function. For each pair of nodes, the decision maker sets a price, which induces a demand that needs to be served. The network may have multiple paths connecting an origin and a destination, so the decision maker chooses how to distribute a demand flow over the different paths, while taking into account the capacity available on the arcs. The goal is to maximize the overall profit for the system, which is determined by the revenues obtained from serving demand minus the costs incurred by transporting demand through the network.

The problem described above has several applications in logistics and freight transportation. For example, consider a package delivery carrier such as FedEx or UPS that operates an expansive service network. The carrier charges prices for packages based on their origins, destinations, and service classes (e.g., next morning, next day, two-day, etc.). The prices will determine the size of demand to be served, since customers are price-sensitive. Because the carrier operates a dense service network, there are possibly multiple options for sending packages from their origins to their destinations. Customers do not necessarily care about which routes their packages take, as long as they are delivered on time. So the carrier can select routes for packages based on transportation costs and capacities. Pricing and routing decisions should ideally be coordinated and made in an integrated manner, since both decisions affect the carrier's profit.

We propose two algorithms for solving the integrated pricing and routing problem. Both algorithms iteratively solves a sub-problem that can be formulated as a minimum cost multicommodity flow (MCMCF) problem. The first algorithm is based on the Frank-Wolfe algorithm (Frank and Wolfe, 1956) (also known as the conditional gradient method (Levitin and Polyak, 1966)). We show that the rate of convergence of this algorithm is $O(1/T)$ where T is the number of iterations. To obtain this result, the algorithm requires the revenue function to be smooth and concave. The second algorithm is a primal-dual algorithm that updates pricing and routing decisions iteratively. In each iteration, the algorithm executes two steps: the first step is to set prices to maximize profit that adjusted by dual variables; the second step is to determine how to split such demands among multiple paths to minimize cost, as well as updating the dual variables. When the objective function is strongly concave, we show that this primal-dual algorithm has a rate of convergence of $O(\log T/T)$ where T is the number of iterations. An advantage of the primal-dual algorithm over the Frank-Wolfe algorithm is that it allows a non-smooth revenue function (e.g., a piece-wise linear demand function).

Our numerical experiments, using randomly generated instances, show that joint pricing and routing can improve profit by more than 10% compared to making pricing and routing decisions separately. A more in-depth analysis of the resulting solutions shows that in portions of the service network where several demands compete for transport capacity, prices are adjusted to free up capacity for demand with a high profit margin. That is, the price of demand with a low profit margin is raised to reduce that demand and reduce the capacity required to serve it, and the price of demand with a high profit margin is dropped to increase that demand and use the capacity that has become available to serve it. The intricate interactions between price adjustments and routing choices that are revealed demonstrate that optimization techniques are necessary to identify and exploit these opportunities to the fullest.

The remainder of the chapter is organized as follows. In Section 3.2, we review relevant literature. In Section 3.3, we introduce the integrated pricing and routing problem and give mathematical formulations. In Section 3.4, we present two algorithms for the solution of the integrated pricing and routing algorithm. In Section 3.5, we discuss the result of our computational experiments. In Section 3.6, we offer some final thoughts.

3.2 Literature Review

Despite the fact that there are numerous papers on network routing problems in which demand is exogenous (e.g., Kennington, 1978), there are only a few papers which consider integrated pricing and routing decisions. Mitra, Ramakrishnan, and Wang (2001) studied a joint pricing and routing problem for revenue maximization in a multi-service network, assuming the set of possible routes is given. Sharkey (2011) studied an integrated pricing and routing problem for a single product, and showed that the problem can be reformulated as a minimum convex cost network flow problem. Lin, Lin, and Young (2009) and Lin and Lee (2015) formulate joint pricing and routing problems for less-than-truckload transportation as an integer concave programming problem.

The routing problem, i.e., finding origin-destination paths for commodities in a network, is a critical component of Service Network Design (SND) problems. SND integrates capacity and routing decisions, i.e., decisions regarding where and how much capacity to install on the links in a network and decisions regarding the paths that commodities follow from their origin to their destination in the network. For a review of SND, we refer the reader to Crainic, 2000 and Wieberneit, 2008.

Our proposed methods use the Frank-Wolfe algorithm (Frank and Wolfe, 1956; Levitin and Polyak, 1966) and the online gradient descent algorithm (Zinkevich, 2003). the Frank-Wolfe algorithm under different assumptions on the objective function and feasible set. Zinkevich (2003) shows that the online gradient descent algorithm achieves $O(\sqrt{T})$ regret for a general convex function when using step size $1/\sqrt{t}$ at iteration t. Hazan, 2016 provide a stronger regret bound, $O(\log T)$, when running with step size $1/\mu t$ at iteration t for a μ-strongly convex objective function.

Another core component of our methods is the solution of a minimum (linear) cost multicommodity flow (MCMCF) problem. The MCMCF problems can be found in a wide range of domains, for example, in transportation and logistics (Krile, 2004), in communication networks (Resende and Pardalos, 2006, Chapter 10) and in scheduling (Vaidyanathan, Jha, and Ahuja, 2007). Even though the MCMCF problem can be formulated as a linear programming and thus solved by a general linear programming solver, the size of instances makes this often inefficient or impractical. As a result, many custom solution methods have developed for solving the MCMCF problem. Comprehensive surveys by Assad (1978) and Kennington (1978) describe many such solution techniques. In addition, a detailed survey of solution approaches for minimum convex cost multicommodity flow problems can be found in Ouorou, Mahey, and Vial (2000).

3.3 Problem Description

We study an integrated pricing and routing problem on a fixed network represented by a directed graph $\mathscr{G} = (\mathscr{N}, \mathscr{E})$, where the set of nodes, $\mathscr{N}$, represents a set of cities and the set of arcs, $\mathscr{E}$, represents connections between pairs of cities. Each arc $e \in \mathscr{E}$ has a per-unit traversal cost c_e and a capacity k_e.

The set of origin-destination (OD) cities for which delivery service is offered is denoted by $\mathscr{J} \subseteq \mathscr{N} \times \mathscr{N}$. We assume that demand is a function of price for each OD pair $j \in \mathscr{J}$

and that this demand function, $d_j(p)$, is known. We also assume that $d_j(p)$ is strictly decreasing, which implies that there is a one-to-one mapping between demands and prices. Let $p_j(d)$ be the inverse demand function of $d_j(p)$, which allows us to recover the price for a given demand. Let $r_j(d) := dp_j(d)$ denote the revenue function, the revenue collected for OD pair $j \in \mathscr{J}$ for demand d. We assume that each revenue function r_j is concave.

The demand of an OD pair $j \in \mathscr{J}$ must be delivered via one or more paths in the network connecting the origin and destination (respecting the capacities of the arcs in the network). We let $\mathscr{P}_j$ denote the set of possible (directed) paths connecting the origin and destination of OD pair $j \in \mathscr{J}$.

The decision maker wants to maximize the profit, i.e., the total revenue collected minus the total delivery cost incurred. The decision maker can set the price (or, equivalently, demand) and choose the delivery routes for each OD pair, where the arc flows induced by the delivery routes cannot exceed the capacity of the arcs.

Let z_j be the demand of OD pair j. We use x_p to represent the demand delivered using path p and v_e to represent the demand carried on arc e. We let arc-path indicator $\delta_e(p)$ equal to 1 if arc e is in path p and 0 otherwise. We write v, x and z to denote the vectors of v_e, x_p and z_j, respectively. A path-based formulation for the integrated pricing and routing problem is

$$\max_{v,x,z} \quad \sum_{j\in\mathscr{J}} r_j(z_j) - \sum_{e\in\mathscr{E}} c_e v_e \tag{3.1a}$$

$$\text{s.t.} \quad z_j = \sum_{p\in\mathscr{P}_j} x_p, \qquad \forall j \in \mathscr{J}, \tag{3.1b}$$

$$v_e = \sum_{j\in\mathscr{J}} \sum_{p\in\mathscr{P}_j} \delta_e(p) x_p, \quad \forall e \in \mathscr{E}, \tag{3.1c}$$

$$v_e \leq k_e, \qquad \forall e \in \mathscr{E}, \tag{3.1d}$$

$$x_p \geq 0, \qquad \forall p \in \mathscr{P}_j, j \in \mathscr{J}. \tag{3.1e}$$

Constraints (3.1b) ensure that the demand for an OD pair is delivered using one or more feasible paths for that demand. Constraints (3.1c) determine the flow on an arc in the network. Constraints (3.1d) ensure that the flow on an arc does not exceed the arc's capacity. Constraints (3.1e) ensure that the demand allocated to a path is non-negative. This path-based formulation has $|\mathscr{J}|+\sum_{j\in\mathscr{J}}|\mathscr{P}_j|+|\mathscr{E}|$ decision variables and $|\mathscr{J}|+2|\mathscr{E}|$ constraints. Note that the number of paths, $|\mathscr{P}|=\sum_{j\in\mathscr{J}}|\mathscr{P}_j|$, is typically large and grows exponentially in the size of network. As the total number of paths becomes prohibitively large even for moderate-size networks, customized solution methods are needed for its solution.

The problem has an equivalent arc-based formulation, which can be found in Appendix B.1, but our proposed algorithms use the path-based formulation because it is more flexible and can incorporate additional constraints on paths arising in real-world applications. For example, suppose we want to impose constraints on feasible prices such that the prices in the solution do not deviate too much from the current prices. If the new price of OD pair j is restricted to the interval $[\underline{p}_j, \bar{p}_j]$, then, because of the one-to-one mapping between price and demand, the demand z_j has to lie in the interval $[\underline{d}_j, \bar{d}_j]$, where $\underline{d}_j$ and $\bar{d}_j$ are the value of the inverse demand function at $\bar{p}_j$ and $\underline{p}_j$, respectively. More importantly, the path-based formulation can be solved efficiently by the column generation technique (Ford and Fulkerson, 1958).

3.4 Algorithms

In this section, we propose two algorithms to solve the pricing and routing problem. Before we describe these algorithms, we review some definitions for convex functions. We say $\nabla f(x)$ is a subgradient of f at x if for any x_1, it holds that

$$f(x_1) \geq f(x) + \nabla f(x)^\top (x_1 - x).$$

The set of subgradients of f at the point x is called the subdifferential of f at x, and is denoted $\partial f(x)$. If a function f is convex and differentiable, then a subgradient is unique at any x and equal to its gradient at x. A subgradient can exist even when f is not differentiable at x. Note that a subgradient of convex function f always exists. We write $\|x\|$ for the euclidean norm of x, i.e., $\|x\| = \sqrt{x^\top x}$. We say a function f is μ-strongly convex if, for any x_1 and x_2, it holds that

$$f(x_2) \geq f(x_1) + \nabla f(x_1)^\top (x_2 - x_1) + \frac{\mu}{2}\|x_2 - x_1\|^2.$$

Recall that $\nabla f(x)$ a subgradient of f at x and $\partial f(x)$ the subdifferential (i.e., the set of subgradients) of f at x. We say a function g is μ-strongly concave if $-g$ is μ-strongly convex. We say a continuous differentiable function f is β-smooth if

$$f(x_2) \leq f(x_1) + \nabla f(x_1)^\top (x_2 - x_1) + \frac{\beta}{2}\|x_2 - x_1\|^2,$$

which is equivalent to say that a continuous differentiable function f has a β-Lipschitz continuous gradient, i.e.,

$$\|\nabla f(x_1) - \nabla f(x_2)\| \leq \beta \|x_1 - x_2\|.$$

3.4.1 Frank-Wolfe Algorithm with Column Generation (FW-CG)

The Frank-Wolfe algorithm (Frank and Wolfe, 1956), also known as the conditional gradient method (Levitin and Polyak, 1966), is a popular first-order method for smooth constrained convex optimization problem of the form $\max\{f(x) \mid x \in X\}$, where f is a concave and smooth function and X is a polyhedron.

At each iteration t, the Frank-Wolfe algorithm evaluates the gradient of function f at current solution x_t, i.e., $\nabla f(x_t)$. Then, it solves a linear programming subproblem given by $w_t = \arg\max_{w \in X} \nabla f(x_t)^\top w$ before updating the solution suing the solution obtained from the linear programming subproblem as $x_{t+1} = x_t + \gamma_t (w_t - x_t)$ where γ_t is step size. The formal description can be found in Algorithm 5. Notice that the Frank-Wolfe algorithm does

Algorithm 5 The Frank-Wolfe (FW) Algorithm.

Input: $x_0 \in X, \gamma_t = 2/(t+2)$ and T
for $t = 0, 1, 2, \ldots, T$ **do**
  Set $w_t \leftarrow \arg\max_{w \in X} \nabla f(x_t)^\top w$
  Set $x_{t+1} \leftarrow x_t + \gamma_t (w_t - x_t)$
Output: x_T

not require a projection onto the feasible set but only depends on solving a linear optimization problem over the constrained set. That is, the Frank-Wolfe algorithm performs well if it is inexpensive to solve the linear programming subproblem. The rate of convergence of the Frank-Wolfe algorithm described in Algorithm 5 in the $\beta-$smooth concave objective function is $O(1/T)$ (Lemma 13 in Appendix B.3, Jaggi, 2013, Theorem 1).

When we apply the Frank-Wolfe algorithm to the integrated pricing and routing problem, we solve a convex optimization problem by solving a sequence of linear optimization problems. It is easy to verify that the linear programming subproblem becomes minimum-cost multicommodity flow problem (MCMCF). One way to efficiently solve this linear programming problem is to apply column generation method to path-flow formulation of MCMCF (Ahuja, Magnanti, and Orlin, 1993).

Column generation can be used to solve LP with a large number of decision variables. Ford and Fulkerson (1958) outline the idea of using column generation form a maximal multicommodity flow problem, where MCMCF is one of its variation, as follows. The problem is solved by considering two problems: the master problem which includes only a subset of decision variables and the pricing problem which determines if a new variables should be included in the master problem to improve objective function.

Recall that the integrated pricing and routing problem in path-based formulation can be written as

$$\begin{aligned}
\max_{x} \quad & \sum_{j\in\mathscr{J}} r_j(\sum_{p\in\mathscr{P}_j} x_p) - \sum_{e\in\mathscr{E}} c_e \sum_{j\in J}\sum_{p\in\mathscr{P}_j} \delta_e(p) x_p \\
\text{s.t.} \quad & \sum_{j\in J}\sum_{p\in\mathscr{P}_j} \delta_e(p) x_p \le k_e, \quad \forall e\in\mathscr{E}, \\
& x_p \ge 0, \quad \forall p\in\mathscr{P}_j, j\in\mathscr{J}.
\end{aligned} \tag{3.2}$$

The objective function of (3.2) is defined as $f(x)$, and a feasible set of (3.2) is defined as X To apply the Frank-Wolfe algorithm (see Algorithm 5), we first need to compute the gradient of $f(x)$. Each element of the gradient $\nabla f(x)$ is given by, for path $p\in\mathscr{P}_j$,

$$\frac{\partial}{\partial x_p} f(x) = r'_j(\sum_{p\in\mathscr{P}_j} x_p) - \sum_{e\in\mathscr{E}}\sum_{j\in J}\sum_{p\in\mathscr{P}_j} \delta_p(e) c_e.$$

Each iteration, the Frank-Wolfe algorithm solves the following linear programming subproblem.

$$\begin{aligned}
\max_{w} \quad & \sum_{j\in\mathscr{J}}\sum_{p\in\mathscr{P}_j} \left[r'_j(\sum_{p\in\mathscr{P}_j} x_{p,t}) - \sum_{e\in\mathscr{E}}\sum_{j\in J}\sum_{p\in\mathscr{P}_j} \delta_p(e) c_e \right] w_p \\
\text{s.t.} \quad & \sum_{j\in J}\sum_{p\in\mathscr{P}_j} \delta_p(e) w_p \le k_e, \quad \forall e\in\mathscr{E}, \\
& w_p \ge 0, \quad \forall p\in\mathscr{P}_j, j\in\mathscr{J},
\end{aligned} \tag{3.3}$$

where $x_{p,t}$ is the demand delivered on path p obtained in the current iteration t from the Frank-Wolfe algorithm. This linear optimization subproblem can be viewed as the minimum-cost multicommodity flow problem (MCMCF) in path-flow formulation with per-unit cost

$$c_{p,t} = \sum_{e\in\mathscr{E}}\sum_{j\in J}\sum_{p\in\mathscr{P}_j} \delta_e(p) c_e - r'_j(\sum_{p\in\mathscr{P}_j} x_{p,t})$$

on path $p\in\mathscr{P}_j$. It can be observed that the number of decision variables which equals to the number of paths is exponentially large in the network size. To solve this linear programming subproblem, we apply column generation method. The master problem is defined similar to (3.3) but contains only a subset of paths (decision variables). A new variable is added to the master problem if it can improve the objective function. For maximization problem, the objective function might increase when a new variable with positive reduced cost is included. Therefore, a new variable is included to the master problem when reduced cost is positive. Let $y_{e,t}$ be a dual variable associated with a capacity constraint on arc e at iteration t. Reduced cost of path $p \in \mathscr{P}_j$ can be computed by

$$r'_j(\sum_{p \in \mathscr{P}_j} x_{p,t}) - \sum_{e \in \mathscr{E}} \delta_e(p)(c_e + y_{e,t}).$$

That is, the pricing problem is to check if there exists a path which has positive reduced cost, i.e.,

$$r'_j(\sum_{p \in \mathscr{P}_j} x_{p,t}) - \sum_{e \in \mathscr{E}} \delta_e(p)(c_e + y_{e,t}) > 0 \iff \sum_{e \in \mathscr{E}} \delta_e(p)(c_e + y_{e,t}) < r'_j(\sum_{p \in \mathscr{P}_j} x_{p,t}). \quad (3.4)$$

If there exists a path which satisfies the condition in (3.4), we include it to the master problem. Otherwise, the solution is optimal. Such path is easy to find by running shortest path algorithm, e.g. Dijkstra's algorithm (Dijkstra, 1959), with modified cost $c_e + y_{e,t}$ on each arc e for each OD pair $j \in \mathscr{J}$ at iteration t. Notice that by apply column generation method all possible paths need not be generated beforehand; instead, path will be generated only as required. The complete definition of the Frank-Wolfe with column generation algorithm (FW-CG) for solving the integrated pricing and routing problem can be found in Algorithm 6.

Before we formally state a theorem, let us recall the definition of big O notation. For two functions $f(T)$ and $g(T) > 0$, we write $f(T) = O(g(T))$ if there exists a constant M_1 and a constant T_1 such that $f(T) \leq M_1 g(T)$ for all $T \geq T_1$. The convergence rate of Algorithm 6 follows immediately from Lemma 13 in Appendix B.3 and Theorem 1 in Jaggi (2013) and can be stated as follows.

Algorithm 6 The Frank-Wolfe with Column Generation (FW-CG) Algorithm.

Input: $x_0 = 0, y_0 = 0, \gamma_t = 2/(t+2)$ and T

Initialize: For each OD pair $j \in \mathscr{J}$, consider a path with minimum cost, i.e., $\mathscr{L}_j = \arg\min_{p \in \mathscr{P}_j} \sum_{e \in \mathscr{E}} \delta_e(p) c_e$.

for $t = 1, 2, \ldots, T$ **do**

  do

    Set $i \leftarrow 0$

    (w_t, y_t) solves

$$\begin{aligned} \max_{w} \quad & \sum_{j \in \mathscr{J}} \sum_{p \in \mathscr{L}_j} \left[r'_j\Big(\sum_{p \in \mathscr{L}_j} x_{p,t} \Big) - \sum_{e \in \mathscr{E}} \sum_{j \in J} \sum_{p \in \mathscr{L}_j} \delta_e(p) c_e \right] w_p \\ \text{s.t.} \quad & \sum_{j \in J} \sum_{p \in \mathscr{L}_j} \delta_e(p) w_p \le k_e, \quad \forall e \in \mathscr{E} \quad (y_{e,t}), \\ & w_p \ge 0, \quad \forall p \in \mathscr{L}_j, j \in \mathscr{J}, \end{aligned}$$

    for OD pairs $j \in \mathscr{J}$ **do**

      if $\min_{p \in \mathscr{P}_j} \sum_{e \in \mathscr{E}} \delta_e^p (c_e + y_{e.t}) < r'_j(\sum_{p \in \mathscr{L}_j} x_t^p)$ **then**

        Set $\mathscr{L}_j \leftarrow \mathscr{L}_j \cup \arg\min_{p \in \mathscr{P}_j} \sum_{e \in \mathscr{E}} \delta_e^p (c_e + y_{e.t})$

        Set $i \leftarrow i + 1$

  while $i > 0$

  Set $x_{t+1} \leftarrow x_t + \gamma_t (w_t - x_t)$

Output: x_T

Theorem 2. *When revenue function is smooth and concave, under Algorithm 6, we have*

$$f(x^*) - f(x_T) \leq O\left(\frac{1}{T}\right).$$

3.4.2 Primal-Dual Algorithm

Next, we propose a Primal-Dual (PD) algorithm. This algorithm uses an iterative method based on applying online learning technique on the Lagrangian relaxation of the original integrated pricing and routing problem. The capacity constraints, which are the only bundle constraint in (3.2), are removed from the constrained set and placed into the objective function. In each iteration, the algorithm has two steps which can easily be solved. In the first step, it determines total demand, or equivalently price, for each OD pair (pricing decision). Then, it decides how to deliver such demand through one or more paths (routing decision).

In online convex optimization framework, a decision maker chooses x_t at iteration t before seeing a convex cost function ℓ_t. The cost of $\ell_t(x_t)$ is then realized to the decision maker. Many algorithms is proposed with the aim to minimize the regret, where the regret after T iterations is defined as

$$\sum_{t=1}^{T} \ell_t(x_t) - \min_{x \in X} \sum_{t=1}^{T} \ell_t(x). \tag{3.5}$$

Online Subgradient Descent (Zinkevich, 2003) which will be used in the first step of the proposed algorithm is a well-known and simple online learning algorithm. The formal description of the algorithm can be found in Algorithm 7, where $\Pi_X(u)$ is a projection of u onto a convex set X. For a general convex loss function, Zinkevich (2003) shows that

Algorithm 7 The Online Subgradient Descent Algorithm.

Input: $x_0 \in X, \eta_t$ and T
for $t = 0, 1, 2, \ldots, T$ **do**
  Set $x_t \leftarrow \Pi_X[x_{t-1} - \eta_t \nabla \ell_t(x_t)]$

Algorithm 7 with an appropriate step size has $O(\sqrt{T})$ regret. Hazan (2016) shows that $O(\log T)$ regret is attainable for a strongly convex loss function. This result is formally stated in Lemma 14. Note that the result still holds for a non-smooth convex loss function.

To obtain the Lagrangian dual problem of integrated pricing and routing problem (3.2), we relax the capacity constraints. Let $y_e \geq 0$ be Lagrange multipliers associated with the capacity constraint of arc e. We let y denote the vector of y_e. The Lagrangian function can be written as

$$L(x,y) = \sum_{j\in\mathscr{J}} r_j(\sum_{p\in\mathscr{P}_j} x_p) - \sum_{e\in\mathscr{E}} c_e \sum_{j\in J}\sum_{p\in\mathscr{P}_j} \delta_e(p)x_p + \sum_{e\in\mathscr{E}} y_e(k_e - \sum_{j\in J}\sum_{p\in\mathscr{P}_j} \delta_e(p)x_p). \tag{3.6}$$

The Lagrangian dual problem is given by

$$\min_{y\geq 0}\max_{x\geq 0} L(x,y) = L(x^*,y^*) = \max_{x\geq 0}\min_{y\geq 0} L(x,y), \tag{3.7}$$

where x^* and y^* be the optimal solutions of (3.7). Because a revenue function is assumed to be concave in a demand for each OD pair $j \in \mathscr{J}$, it follows that the Lagrangian function (3.6) is concave in total demand z_j of each OD pair $j \in \mathscr{J}$ as well as concave in flow on path x_p for each $p \in \mathscr{P}_j$. Moreover, we can observe that the Lagrangian function (3.6) is convex (linear) in dual variable y_e for all $e \in \mathscr{E}$.

In the first step of the algorithm, we want to determine the total demand of each OD pair so that the Lagrangian function (3.6) is maximized. Recall that $z_j = \sum_{p\in\mathscr{P}_j} x_p$. Let z be the vector of z_j for all $j \in \mathscr{J}$. Suppose y is constant. We want to find

$$\begin{aligned} z &= \underset{\substack{z\geq 0 \\ z_j=\sum_{p\in\mathscr{P}_j} x_p}}{\arg\max}\ L(x,y) \\ &= \underset{\substack{z\geq 0 \\ z_j=\sum_{p\in\mathscr{P}_j} x_p}}{\arg\max} \sum_{j\in\mathscr{J}} r_j(z_j) - \sum_{e\in\mathscr{E}} (c_e + y_e) \sum_{j\in\mathscr{J}}\sum_{p\in\mathscr{P}_j} \delta_e(p)x_p. \end{aligned}$$

It is obvious that the problem is separable for each OD pair j, i.e., for all $j \in \mathscr{J}$, we have

$$z_j = \underset{\substack{z_j \geq 0 \\ z_j = \sum_{p \in \mathscr{P}_j} x_p}}{\arg\max} \; r_j(z_j) - \sum_{e \in \mathscr{E}} (c_e + y_e) \sum_{p \in \mathscr{P}_j} \delta_e(p) x_p. \tag{3.8}$$

That is, z_j is the solution of the profit maximization for each OD pair $j \in \mathscr{J}$. We observe that to obtain z_j the second term in (3.8) must be minimized. Since the term is basically the modified cost of routing demand of j through paths, all demand must be delivered using only the smallest adjusted cost path. Such path is easily identified by running shortest path algorithm, e.g. Dijkstra's algorithm (Dijkstra, 1959) with modified cost $c_e + y_e$ on each arc. Let p_j^* be the shortest cost path of OD pair j with modified cost $c_e + y_e$, i.e., $p_j^* = \arg\min_{p \in \mathscr{P}_j} \sum_{e \in \mathscr{E}} (c_e + y_e) \delta_e(p)$. Therefore, we can write

$$z_j = \arg\max_{z_j \geq 0} r_j(z_j) - c_{p_j^*} z_j, \tag{3.9}$$

where $c_{p_j^*}$ is the cost of the shortest cost path with modified cost $c_e + y_e$ of OD pair j, i.e., $c_{p_j^*} = \sum_{e \in \mathscr{E}} \delta_e(p_j^*) c_e$. At each iteration the first step applies Online Supergradient Ascent to the maximization problem in (3.9). That is, for each $j \in \mathscr{J}$, z_j moves in the direction of a supergradient of objective function found in (3.9).

There are three main underlying reasons why we choose to operate on the total demand on each OD pair instead of the flow on each possible path. Firstly, we do not require to enumerate all possible paths. Secondly, the number of OD pairs is much smaller than the number of all possible paths, which makes the algorithm computationally efficient. Lastly, the algorithm provides a stronger regret guarantee when r_j is assumed to be strongly concave in z_j for all $j \in \mathscr{J}$. This result follows from the result of Online Supergradient Ascent. However, when r_j is assumed to be strongly concave in z_j, it does not imply that r_j is strongly concave in x_p for all $p \in \mathscr{P}_j$ (see more discussion in Section 3.5.1).

As we mention earlier, the convergence guarantee of Online Supergradient Ascent does not require the Lagrangian function to be smooth in total demand z_j for all $j \in \mathscr{J}$. That is, the result still follows when, for example, price function p_j is piece-wise linear in z_j. This relationship is often observed when there is a price competition.

In the second step of the algorithm, suppose the total demand of each OD pair is fixed from the first step, we want to determine the amount of demand to be delivered on different paths. Based on the Lagrangian dual problem in (3.7), because z is fixed from the first stage we want to solve

$$\min_{y\geq 0} \max_{\substack{x\geq 0 \\ z_j=\sum_{p\in\mathscr{P}_j} x_j}} L(x,y). \tag{3.10}$$

We know that

$$\begin{aligned}
\max_{x:\sum_{p\in\mathscr{P}_j} x_p = z_j} L(x,y) = \max_{x} \quad & \sum_{j\in\mathscr{J}} r_j(\sum_{p\in\mathscr{P}_j} x_p) - \sum_{e\in\mathscr{E}} c_e \sum_{j\in J} \sum_{p\in\mathscr{P}_j} \delta_e(p) x_p \\
& + \sum_{e\in\mathscr{E}} y_e (k_e - \sum_{j\in J} \sum_{p\in\mathscr{P}_j} \delta_e(p) x_p) \\
\text{s.t.} \quad & \sum_{p\in\mathscr{P}_j} x_p = z_j, \quad \forall j \in \mathscr{J}, \\
& x_p \geq 0, \quad \forall p \in \mathscr{P}_j, j \in \mathscr{J},
\end{aligned}$$

and it can be observed that $y_e \geq 0$ for all $e \in \mathscr{E}$ can be viewed as Lagrange multipliers associated with the capacity constraints in the problem in (3.10). Therefore, x and y can be obtained by solving

$$\begin{aligned}
\max_{x} \quad & \sum_{j\in\mathscr{J}} r_j(\sum_{p\in\mathscr{P}_j} x_p) - \sum_{e\in\mathscr{E}} c_e \sum_{j\in J} \sum_{p\in\mathscr{P}_j} \delta_e(p) x_p \\
\text{s.t.} \quad & \sum_{p\in\mathscr{P}_j} x_p = z_j, \quad \forall j \in \mathscr{J}, \\
& \sum_{j\in J} \sum_{p\in\mathscr{P}_j} \delta_e(p) x_p \leq k_e, \quad \forall e \in \mathscr{E} \quad (y_e \geq 0), \\
& x_p \geq 0, \quad \forall p \in \mathscr{P}_j, j \in \mathscr{J}.
\end{aligned}$$

The first term in objective function, $\sum_{j\in\mathscr{J}} r_j(\sum_{p\in\mathscr{P}_j} x_p) = \sum_{j\in J} r_j(z_j)$, is independent of x, so we can ignore this term. Thus, x and y can be obtained from

$$\begin{aligned}
\min_{x} \quad & \sum_{e\in\mathscr{E}} c_e \sum_{j\in J} \sum_{p\in\mathscr{P}_j} \delta_e(p) x_p \\
\text{s.t.} \quad & \sum_{p\in\mathscr{P}_j} x_p = z_j, \qquad \forall j \in \mathscr{J}, \\
& -\sum_{j\in J} \sum_{p\in\mathscr{P}_j} \delta_e(p) x_p \geq -k_e, \quad \forall e \in \mathscr{E} \quad (y_e \geq 0), \\
& x_p \geq 0, \qquad \forall p \in \mathscr{P}_j, j \in \mathscr{J},
\end{aligned} \tag{3.11}$$

which is the minimum-cost multicommodity flow problem (MCMCF) in path-flow formulation. That is, in the second stage the total demand must be spit in the way that minimize the total cost while satisfying the capacity constraints. Observe that the problem (3.11) might be infeasible because the demand constraints which force the total demand of each OD pair to be predetermined z_j might cause the violation in capacity constraints. Therefore, we introduce a dummy arc which is uncapacitated for each OD pair to take care residual demand which cannot satisfy the capacity constraints. A large enough per-unit cost of such arc must be set so that the arc will not be used if not necessary.

We can apply the same technique, column generation method, as described in Section 3.4.1 to solve MCMCF. Specifically, a new variable (path) is added to the master problem when the reduced cost is negative, that is, the pricing problem is to check if for OD pair j there exists a path which satisfies

$$\sum_{e\in\mathscr{E}} \delta_e(p)(c_e + y_e) - \sigma_j < 0 \iff \sum_{e\in\mathscr{E}} \delta_e(p)(c_e + y_e) < \sigma_j, \tag{3.12}$$

where σ_j is a dual variable associated with the demand constraint of OD pair j. Such path can be found easily by running shortest path algorithm such as Dijkstra's algorithm (Dijkstra, 1959) with modified cost $c_e + y_e$ on arc e. Therefore, the MCMCF problem in (3.11) can be solved efficiently by applying column generation method without having to enumerate all possible path. We state the formal description of PD when objective function is strongly concave in Algorithm 8. Note that we write $\max(x,0)$ as $[x]^+$.

Algorithm 8 The Primal-Dual (PD) Algorithm for Integrated Pricing and Routing.

Input: $x_0 = 0, y_0 = 0, z_0 = 0, \mu$ and T
Initialize: For each OD pair $j \in \mathscr{J}$, add an uncapacitated dummy arc with cost $r_j'(0)$ and consider a path with minimum cost, i.e., $\mathscr{L}_j = \arg\min_{p\in\mathscr{P}_j} \sum_{e\in\mathscr{E}} \delta_e(p)c_e$.
for $t = 1,2,\ldots,T$ **do**
  for OD pairs $j \in \mathscr{J}$ **do**
    Update z_t

$$z_{j,t} \leftarrow \left[z_{j,t-1} - \frac{1}{\mu_j t}(\partial - r_j(z_{j,t-1}) + \min_{p\in\mathscr{P}_j} \sum_{e\in\mathscr{E}} (c_e + y_{e,t-1})\delta_e(p)) \right]^+ . \tag{3.13}$$

  do
    Set $i \leftarrow 0$
    (x_t, y_t) solves

$$\begin{aligned} \min_x \quad & \sum_{e\in\mathscr{E}} c_e \sum_{j\in J} \sum_{p\in\mathscr{P}_j} \delta_e(p)x_p \\ \text{s.t.} \quad & \sum_{p\in\mathscr{P}_j} x_p = z_{j,t}, \quad \forall j \in \mathscr{J}, \\ & -\sum_{j\in J}\sum_{p\in\mathscr{P}_j} \delta_e(p)x_p \geq -k_e, \quad \forall e \in \mathscr{E} \quad (y_e \geq 0), \\ & x_p \geq 0, \quad \forall p \in \mathscr{P}_j, j \in \mathscr{J}, \end{aligned}$$

    for OD pairs $j \in \mathscr{J}$ **do**
      if $\min_{p\in\mathscr{P}_j} \sum_{e\in\mathscr{E}} \delta_e(p)(c_e + y_{e,t}) < \sigma_j$ **then**
        Set $\mathscr{L}_j \leftarrow \mathscr{L}_j \cup \arg\min_{p\in\mathscr{P}_j} \sum_{e\in\mathscr{E}} \delta_e(p)(c_e + y_{e,t})$
        Set $i \leftarrow i+1$
  while $i > 0$
Output: $\bar{x} = \sum_{t=1}^T x_t/T$ and $\bar{y} = \sum_{t=1}^T y_t/T$

Theorem 3. *When revenue function is strongly concave in total demand, under Algorithm 8, we have*

$$\left| \frac{1}{T}\sum_{t=1}^T L(x_t,y_t) - L(x^*,y^*) \right| \leq O\left(\frac{\log T}{T}\right).$$

Appendix B.2.1 provides the detailed proof.

Theorem 4. *When revenue function is strongly concave in total demand, under Algorithm 8, we have*

$$\|\bar{z} - z^*\|^2 \leq O\left(\frac{\log T}{T}\right).$$

The proof can be found in Appendix B.2.2.

We remark that unlike the Frank-Wolfe algorithm, the PD algorithm (Algorithm 8) does not require the objective function to be smooth. We obtain the same rates of convergence in Theorem 3 and Theorem 4 for strongly concave and both smooth and non-smooth objective function. Moreover, the PD algorithm can be applied even if the objective function is not strongly concave. In that case, Equation (3.13) in Algorithm 8 needs to be modified. Specifically, for general concave (not necessarily strongly concave) objective function, Equation (3.13) can use a step size of $1/\sqrt{t}$ in place of the step size $1/\mu_j t$. However, for general concave objective, we can only conclude that rate of convergence in term of average regret of this algorithm is $O(1/\sqrt{T})$.

To summarize this section, we proposed two algorithms, Frank-Wolfe with Column Generation (FW-CG) and primal-dual algorithm (PD), for solving integrated pricing and routing problem. When revenue function is smooth and concave, FW-CG achieves the rate of convergence in term of objective function of $O(1/T)$. However, PD still works without a smoothness assumption. We show that for strongly concave revenue function the rates of convergence of PD in term of average regret and total demand are both $O(\log T/T)$. We note that we can alter PD so that the algorithm allows general concave revenue function. Specifically, step size in the first step when updating total demand of each OD pair needs to be modified. At iteration t, modified PD uses step size $1/\sqrt{t}$ instead of $1/\mu_j t$. This modified algorithm with new step size has the rate of convergence in term of average regret of $O(1/\sqrt{T})$ for general concave revenue function. The proof is similar to The proof of Theorem 3, but we use $O(\sqrt{T})$ bound (Zinkevich, 2003) when we apply Online Supergradient Ascent. Note that when revenue function is not strongly concave, we are not able to conclude the rate of convergence in term of total demand.

3.5 Numerical Experiments

In the experiments below, we consider a service network with 10 cities and connections between all cities, i.e., a complete directed graph $\mathscr{G} = (\mathscr{N}, \mathscr{E})$ with $|\mathscr{N}| = 10$ and $|\mathscr{E}| = 90$, resulting in $|\sum_{j \in J} P_j| = 9,864,090$ paths. Let $U\{\underline{u}, \bar{u}\}$ denote the discrete uniform distribution which can have integer values from $\underline{u}$ and $\bar{u}$. The per-unit traversal cost and the capacity are drawn from $U\{1,5\}$ as well for all arcs $e \in \mathscr{E}$, i.e., $c_e \sim U\{1,5\}$ and $k_e \sim U\{1,5\}$.

3.5.1 Demand Models

We consider two different types of demand models.

Linear Demand Model.

We assume linear demand models for all origin-destination (OD) pairs $j \in J$ with different parameter values:

$$d_j(p) = a_j - b_j p,$$

where $a_j > 0$ and $b_j > 0$. Equivalently, we can write the price as a linear function of demand as

$$p_j(d) = \frac{a_j}{b_j} - \frac{1}{b_j} d.$$

Recall that the revenue function $r_j(d) := p_j(d)d$ requires to be concave. For linear demand model, we can then write

$$r_j(d) = \frac{a_j}{b_j}d - \frac{1}{b_j}d^2,$$

which is strongly concave in d, because $r_j''(d) = -2/b_j < 0$. Note that although revenue function of linear demand model is strongly concave in total demand, it is concave but not strongly concave in the flow on each path. To see this, we write total demand as a summation of flows on each possible path. We have $\nabla^2 r_j(d) = -2J_{|\mathscr{P}_j|}$, where J_n is an $n \times n$ matrix of ones. Its eigenvalues are $-2|\mathscr{P}_j|$ and 0, which are less than or equal to zero. Therefore, revenue function of linear demand model is concave but not strongly concave in flow on each path. We assume that the slope parameter of the price function is drawn from $U\{1,5\}$ for all OD pairs $j \in \mathscr{J}$, i.e., $1/b_j \sim U\{1,5\}$.

Piece-wise Linear Demand Model.

We assume that, for each OD pair $j \in J$, the demand function can be written as

$$d_j(p) = \begin{cases} a_{j,1} - b_{j,1}p & p \le p_0 \\ a_{j,2} - b_{j,2}p & p > p_0, \end{cases}$$

where $a_{j,1} > 0, a_{j,2} > 0, b_{j,1} > 0$ and $b_{j,2} > 0$. Equivalently, we can write the price as a piece-wise linear function of demand as

$$p_j(d) = \begin{cases} \frac{a_{j,1}}{b_{j,1}} - \frac{1}{b_{j,1}}d & d \le d_0 \\ \frac{a_{j,2}}{b_{j,2}} - \frac{1}{b_{j,2}}d & d > d_0. \end{cases}$$

Specifically, we are interested in the demand function of the following form.

$$p_j(d) = \begin{cases} \frac{a_j}{b_j} - \frac{1}{b_j}d & d \le \frac{a_j}{4} \\ \frac{5a_j}{4b_j} - \frac{2}{b_j}d & d > \frac{a_j}{4}. \end{cases}$$

The revenue function can then be written as

$$r_j(d) = \begin{cases} \frac{a_j}{b_j}d - \frac{1}{b_j}d^2 & d \le \frac{a_j}{4} \\ \frac{5a_j}{4b_j}d - \frac{2}{b_j}d^2 & d > \frac{a_j}{4}. \end{cases}$$

It can be seen that this revenue function is not differentiable at $a_j/4$. The subdifferential of revenue function is

$$\partial r_j(d) = \begin{cases} \frac{a_j}{b_j} - \frac{2}{b_j}d & d < \frac{a_j}{4} \\ \left[\frac{a_j}{2b_j}, \frac{a_j}{4b_j}\right] & d = \frac{a_j}{4} \\ \frac{a_j}{b_j} - \frac{4}{b_j}d & d > \frac{a_j}{4}, \end{cases}$$

which is decreasing function in d. Therefore, the revenue function is concave. However, because the revenue function is non-smooth, only PD is applicable to this setting. We assume that the slope parameter of the price function is drawn from $U\{1,5\}$ for all OD pairs $j \in \mathscr{J}$, i.e., $1/b_j \sim U\{1,5\}$.

3.5.2 Experimental Results

We test four different distributions, $U\{7,11\}$, $U\{9,13\}$, $U\{11,15\}$, and $U\{13,17\}$ for the intercept parameter of the price function, a_j/b_j. Consequently, the same demand can be induced by setting a higher price, and higher profit is obtained for the same demand. We generate three instances for each distribution of the intercept parameter of price function while assuming other parameters have the same distributions.

We investigate the benefit of integrating pricing and routing decisions by considering three different settings:

1. PRICING ONLY: The decision maker sets optimal prices for each OD pair given that the demand for the OD pair only uses the path consisting of the direct arc that links the origin to the destination.

2. PRICING $\rightarrow$ ROUTING: The decision maker sets optimal prices for each OD pair assuming that the demand for the OD pair only uses the path consisting of the direct arc that links the origin to the destination, but after setting the prices, the decision maker optimally chooses one or more routes for each OD pair to serve the resulting demand.

3. PRICING + ROUTING: The decision maker simultaneously and optimally determines the prices and the routes to serve the resulting demand for each OD pair, i.e., solve the integrated pricing and routing problem.

Linear Demand Model.

Figure 3.1 shows the average profit for the three settings when the intercept term of the price function, a_j/b_j, is drawn from $U\{7,11\}$, $U\{9,13\}$, $U\{11,15\}$, and $U\{13,17\}$. We observe a slight increase in profit, 5.23%, 3.17%, 1.87%, and 1.08%, respectively, for PRICING $\rightarrow$ ROUTING over PRICING ONLY, and a significant increase in profit, 11.39%, 10.96%, 11.08%, and 11.54%, respectively, for PRICING + ROUTING over PRICING ONLY.

We also see that the improvement in average profit of PRICING $\rightarrow$ ROUTING over PRICING ONLY decreases as the intercept term of price function increases. The reason is that the optimal demand for each OD pair never decreases (and likely increases) when the intercept term of price function increases. As a result, the capacity utilization increases which restricts the allocation of demand across the different paths when deciding the routes

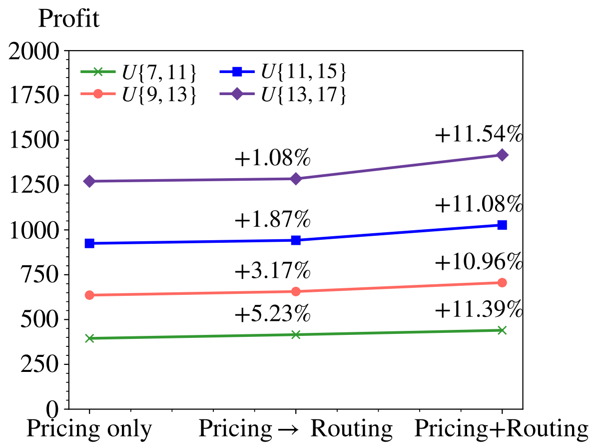


Figure 3.1: The average profit of PRICING ONLY, Pricing $\rightarrow$ Routing and PRICING + ROUTING when the intercept term of the price function, a_j/b_j, is drawn from $U\{7,11\}$, $U\{9,13\}$, $U\{11,15\}$ and $U\{13,17\}$.

for the demands.

Figure 3.2 shows the average arc capacity utilization in the PRICING ONLY, PRICING $\rightarrow$ ROUTING, and PRICING + ROUTING solutions when the intercept term of the price function, a_j/b_j, is drawn from $U\{7,11\}$, $U\{9,13\}$, $U\{11,15\}$, and $U\{13,17\}$. This graph shows that effectively using the capacity in the service network is critical to achieving high profits. Even though the demand is the same in the PRICING ONLY and PRICING $\rightarrow$ ROUTING settings, the arc capacity utilization increases because some of the demand is allocated to longer, but cheaper, alternative paths.

Figure 3.1 clearly shows that simultaneously optimizing pricing and routing decisions pays off. In the PRICING + ROUTING solution, the price for an OD pair often differs from the price for the same OD pair in the PRICING ONLY solution to induce more demand for more profitable OD pairs and less demand to less profitable OD pairs. Two factors impact the profit that can be achieved when simultaneously optimizing pricing and routing decisions: (1) the number profitable paths, i.e., paths for which the per-unit revenue is greater than the per-unit cost, and (2) the capacity utilization. As mentioned earlier, an increase in the intercept term of the price function increases the capacity utilization, which,

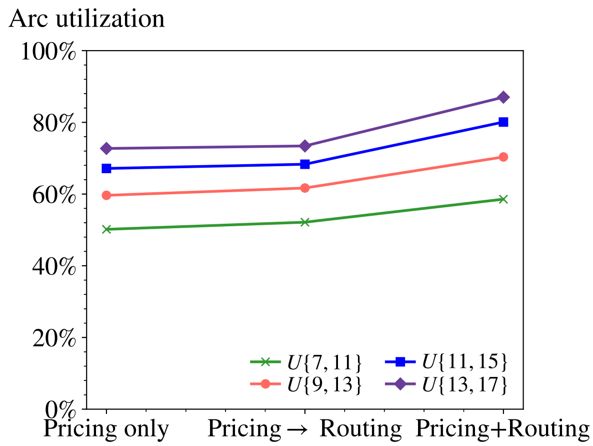


Figure 3.2: The average capacity utilization of PRICING ONLY, PRICING → ROUTING and PRICING + ROUTING when demand function is linear, and the intercept term of the price function, a_j/b_j, is drawn from $U\{7,11\}$, $U\{9,13\}$, $U\{11,15\}$ and $U\{13,17\}$.

in turn, reduces the routing flexibility. The larger the number of profitable paths for an OD pair, the more options can be exploited by the optimization.

Figure 3.3 shows the average number of paths for which the per-unit revenue is greater than the per-unit cost when using the prices set in the PRICING ONLY solution, when the intercept term of price function a_j/b_j is drawn from $U\{7,11\}$, $U\{9,13\}$, $U\{11,15\}$, and $U\{13,17\}$. We see that the average number of profitable paths increases exponentially when the intercept term of the price function increases.

The interaction between these two factors has different effects in different instances, which explains why we do not see a monotonic change in average profit when the intercept term of the price function is increased.

Next, we investigate the solution to a single instance in more depth (the intercept term of price function for this instance is drawn from $U\{11,15\}$). We focus on portion of the service network consisting of four nodes and five arcs; the arcs with their cost and capacity are shown in Figure 3.4.

In Table 3.1, we show, for each of the five OD pairs corresponding to the five arcs, the size of the demand, along how many paths the demand is routed, and the profit obtained

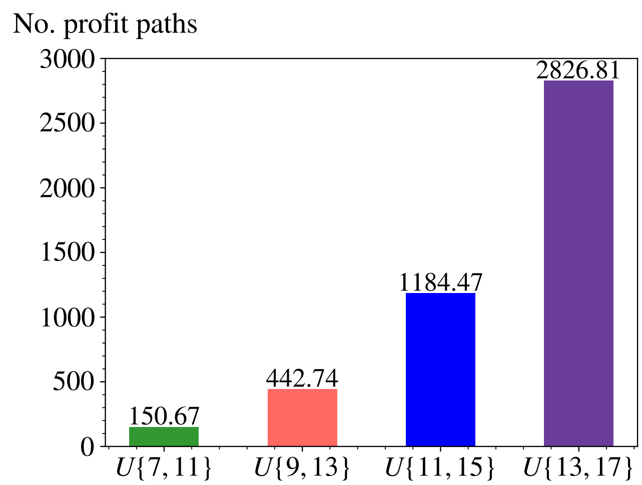


Figure 3.3: The average number of profitable paths when using the prices of the PRICING ONLY solution when demand function is linear, and the intercept term of price function a_j/b_j is drawn from $U\{7,11\}$, $U\{9,13\}$, $U\{11,15\}$ and $U\{13,17\}$.

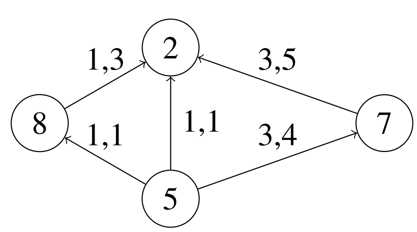


Figure 3.4: Cost (first element) and capacity (second element) of a subset of arcs in the network

from serving the demand in the optimal solution for the PRICING ONLY and for the PRICING + ROUTING setting. Furthermore, in Table 3.2, we show, for each of the arcs, the total demand on the arc and the arc's capacity utilization in the optimal solution for the PRICING ONLY and for the PRICING + ROUTING setting as well as the specific demands allocated to the arc and their contribution to the arc's capacity utilization.

Table 3.1: Demands of a subset of OD pairs in PRICING ONLY and PRICING + ROUTING

OD	PRICING ONLY		
	price	demand	total profit
(5,2)	10.00	1.00	9.00
(5,7)	8.00	1.00	5.00
(5,8)	6.00	1.00	5.00
(7,2)	9.00	1.50	9.00
(8,2)	7.00	1.50	9.00

OD	PRICING + ROUTING				
	price	demand	#routes	avg per-unit profit	total profit
(5,2)	9.22	1.39	3	7.87	10.93
(5,7)	8.00	1.00	1	5.00	5.00
(5,8)	7.82	0.64	1	6.82	4.33
(7,2)	9.22	1.44	1	6.22	8.99
(8,2)	7.40	1.40	1	6.40	8.96

We see in Table 3.1 that the demand for an OD pair in the solution for the PRICING + ROUTING setting can go up or down when compared to the demand in the solution for the PRICING ONLY setting. Because in the PRICING + ROUTING setting, the demand can be delivered using multiple paths, the OD pairs are competing for the capacity on the arcs. Therefore, in an optimal solution, the capacity should be used by the more profitable OD pairs. In fact, for OD pairs with a high per-unit profit, it may be beneficial to induce more demand, by setting a lower price, while for OD pairs with low per-unit profit, it may be beneficial to induce less demand, by setting a higher price, in order to create capacity that can be exploited by more profitable OD pairs. For example, in the solution for the PRICING + ROUTING setting, the demand for OD pair (5,2), which has a high per-unit profit, is increased and delivered using the paths $5 \rightarrow 8 \rightarrow 2$ and $5 \rightarrow 7 \rightarrow 2$. The demands

Table 3.2: Flows and utilization of a subset of arcs in PRICING ONLY and PRICING + ROUTING

arc	PRICING ONLY			PRICING + ROUTING		
	OD	flow	utilization	OD	flow	utilization
5→2		1.00	1.00		1.00	1.00
	(5,2)	1.00	1.00	(5,2)	1.00	1.00
5→7		1.00	0.25		1.02	0.26
	(5,7)	1.00	0.25	(5,2)	0.02	0.01
				(5,7)	1.00	0.25
5→8		1.00	1.00		1.00	1.00
	(5,8)	1.00	1.00	(5,2)	0.36	0.36
				(5,8)	0.64	0.64
7→2		1.50	0.30		5.00	1.00
	(7,2)	1.50	0.30	(5,2)	0.02	0.00
				(7,2)	1.44	0.29
				(7,0)	*0.25*	*0.05*
				(0,2)	*0.46*	*0.09*
				(7,8)	*2.82*	*0.56*
8→2		1.50	0.50		3.00	1.00
	(8,2)	1.50	0.50	(5,2)	0.36	0.12
				(8,2)	1.40	0.47
				(0,2)	*0.49*	*0.16*
				(1,2)	*0.52*	*0.17*
				(8,1)	*0.23*	*0.08*

of OD pairs (5,8), (5,7), and (7,2), which have a low per-unit profit, are decreased to free up capacity for the demand of OD pair (5,2).

In the PRICING + ROUTING setting, where demand can be satisfied using multiple paths, the arc capacity has to be shared by OD pairs. Because different OD pairs have different price functions, it becomes virtually impossible to construct solutions with high profit solutions without the use of optimization. To illustrate, the demand of OD pair (5,2) is 1.0 in the solution to the PRICING ONLY setting, which uses the entire capacity of arc (5,2). If the arc (5,2) would have been uncapacitated, the optimal demand for OD pair (5,2) in the PRICING ONLY setting would have been 2.75. In the PRICING + ROUTING setting, the optimal demand for OD pair (5,2) is no longer restricted (only) by the capacity of arc (5,2). Therefore, the price for OD pair (5,2) is reduced from 10.00 to 9.22 to induce a higher demand of 1.39. This is still much less than 2.75, because (1) the costs of the alternative paths, and (2) the competition for capacity on the arcs on the alternative paths. In the solution for the the PRICING + ROUTING setting, the demand of OD pair (5,2) is split among paths 5→2 (1 unit), 5→8→2 (0.36 units), and 5→7→2 (0.02 units). The per-unit cost of path 5→2 is 1.0, while the per-unit costs of the paths 5→8→2 and 5→7→2 are 2 and 6, respectively; these paths are clearly more expensive. Since 0.36 of the capacity of arc 5→8 is allocated to the demand of OD pair (5,2), the demand of OD pair (5,8), which in the solution for PRICING ONLY setting used the entire capacity of arc $5 \rightarrow 8$, is decreased to 0.64 by increasing its price from 6.00 to 7.82.

Piece-wise Linear Demand Model.

Figure 3.5 shows the average profit for the three settings when the demand function is piece-wise linear and the intercept term of the price function, a_j/b_j, is drawn from $U\{7,11\}$, $U\{9,13\}$, $U\{11,15\}$, and $U\{13,17\}$. Similar to linear demand case, we observe a slight increase in profit, 6.62%, 4.36%, 2.86%, and 2.01%, respectively, for PRICING → ROUTING over PRICING ONLY, and a significant increase in profit, 12.95%, 12.79%, 13.41%, and 14.38%, respectively, for PRICING + ROUTING over PRICING ONLY.

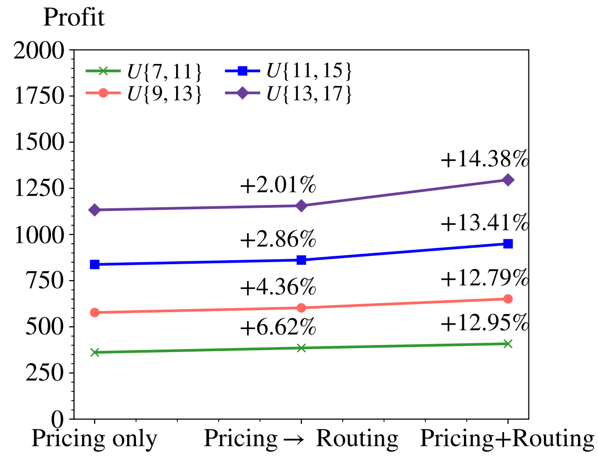


Figure 3.5: The average profit of PRICING ONLY, Pricing → Routing and PRICING + ROUTING when demand function is piece-wise linear and the intercept term of the price function, a_j/b_j, is drawn from $U\{7,11\}$, $U\{9,13\}$, $U\{11,15\}$ and $U\{13,17\}$.

Figure 3.6 shows the average arc capacity utilization in the PRICING ONLY, PRICING → ROUTING, and PRICING + ROUTING solutions when the demand function is piece-wise linear and the intercept term of the price function, a_j/b_j, is drawn from $U\{7,11\}$, $U\{9,13\}$, $U\{11,15\}$, and $U\{13,17\}$. Figure 3.7 shows the average number of paths for which the

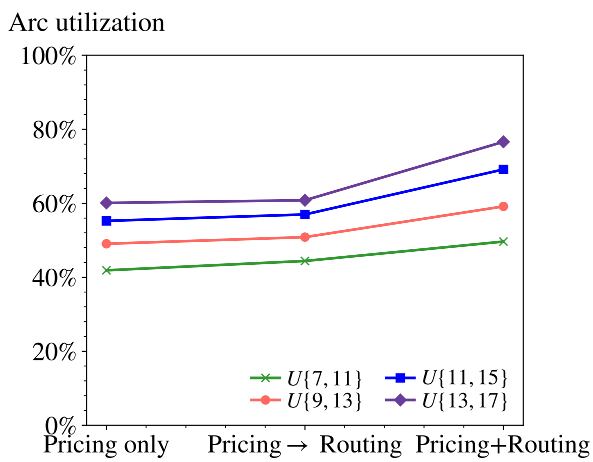


Figure 3.6: The average capacity utilization of PRICING ONLY, PRICING → ROUTING and PRICING + ROUTING when demand function is piece-wise linear and the intercept term of the price function, a_j/b_j, is drawn from $U\{7,11\}$, $U\{9,13\}$, $U\{11,15\}$ and $U\{13,17\}$.

per-unit revenue is greater than the per-unit cost when using the prices set in the PRICING

ONLY solution, when the demand function is piece-wise linear and the intercept term of price function a_j/b_j is drawn from $U\{7,11\}$, $U\{9,13\}$, $U\{11,15\}$, and $U\{13,17\}$.

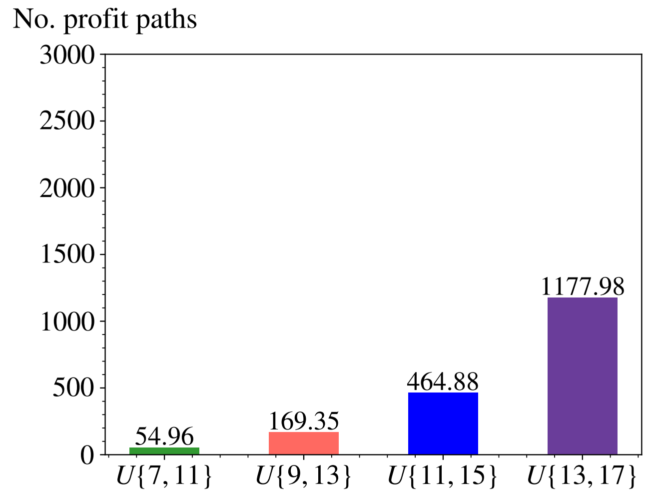


Figure 3.7: The average number of profitable paths when using the prices of the PRICING ONLY solution when the intercept term of price function a_j/b_j is drawn from $U\{7,11\}$, $U\{9,13\}$, $U\{11,15\}$ and $U\{13,17\}$.

As discussed in Section 3.5.2, because of limited arc capacity, we observe a monotonic decrease in average profit improvement of PRICING $\rightarrow$ ROUTING over PRICING ONLY when the intercept term of the price function increases. However, because of the interaction between the capacity utilization and the number of profitable paths, we cannot observe such trend in average profit improvement of PRICING + ROUTING over PRICING ONLY.

3.5.3 Computation Time and Convergence Rates.

In this section, we will study the convergence rate to the optimal solution of the Frank-Wolfe algorithm with column generation (FW-CG) and the Primal-Dual algorithm (PD) when applicable. We test one problem instance with the intercept term of the price function a_j/b_j drawn from $U\{11,15\}$.

We apply the proposed algorithms, when applicable, to the simulated network for 10000 iterations. We measure performance of the algorithms using the metrics found in Table 3.3. The first metric, log loss, measures the gap between objective value obtained from proposed algorithm and optimal objective value. The second metric, log optimal demand gap measures the total difference between demand solutions obtained from proposed algorithm and optimal demands. The third metric, log optimal path flow gap measures the total difference between path flow solutions obtained from proposed algorithm and optimal path flows, and the last metric, log optimal arc flows gap, measures the total difference between arc flow solutions obtained from proposed algorithm and optimal arc flows.

Table 3.3: The performance metrics.

Metric	FW-CG	PD
log loss	$\log(f(x^*) - f(x_t))$	$\log\lvert\frac{1}{t}\sum_{s=1}^{t} L(x_s, y_s) - L(x^*, y^*)\rvert$
log optimal demand gap	$\log\lVert z_t - z^*\rVert^2$	$\log\lVert\bar{z}_t - z^*\rVert^2$
log optimal path flow gap	$\log\lVert x_t - x^*\rVert^2$	$\log\lVert\bar{x}_t - x^*\rVert^2$
log optimal arc flows gap	$\log\lVert v_t - v^*\rVert^2$	$\log\lVert\bar{v}_t - v^*\rVert^2$

Linear Demand Model.

Recall that we assume that each OD pair has a linear relationship between demand and price. That is, the revenue function is both smooth and strongly concave. Therefore, both the Frank-Wolfe algorithm with column generation (FW-CG) and the Primal-Dual algorithm (PD) can be applied in order to solve the integrated pricing and routing problem.

We compare the convergence results of two algorithms when measured against log iteration across four metrics in Figure 3.8. Note that we will suppress base 10 of logarithmic function and write log to represent $\log_{10}$ in this section.

Figure 3.8 shows that log loss and log iteration has linear relationships with negative slope for both algorithms. This numerical result coincides with the theoretical results in Lemma 13 and Theorem 3. Moreover, we can observe that FW-CG has a faster convergence rate in term of objective value than PD. Note that although FW-CG converges faster, it requires an additional condition, smoothness, on revenue function, while PD only needs revenue function to be strongly concave. Even though we only show the convergence of demand when using PD in Theorem 4, Figure 3.8 show numerically the convergence of

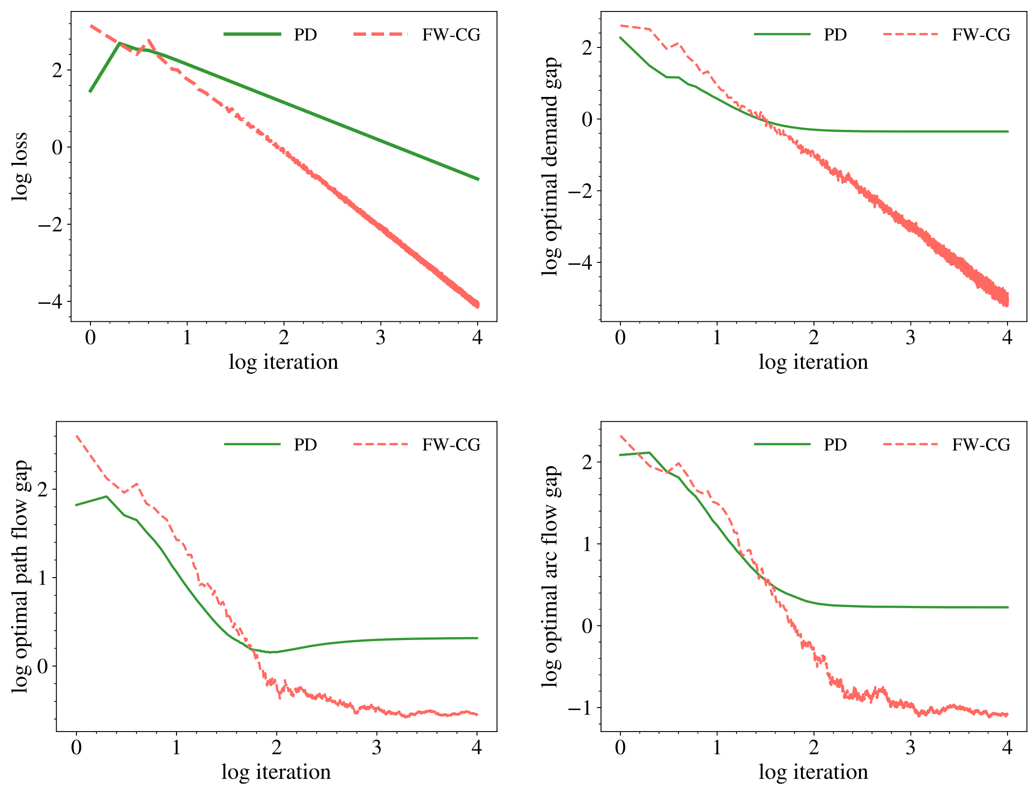


Figure 3.8: The plots compare log loss (upper left), log optimal demand gap (upper right), log optimal path flow gap (lower left) and log optimal arc flows gap (lower right) of PD and FW-CG against log iteration (x-axis) when demand function is linear.

demand, path flow and arc flow from both proposed algorithms. In fact, we can observe that they seem to converge in the same order. When we examine the run time, PD and FW-CG take 96.93 seconds and 67.65 seconds per 10000 iterations respectively.

Next, we show how sensitive of running time to an increase in the number of nodes of both algorithms. Figure 3.9 plots average time per iterations (seconds) used by PD and FW-CG when demand function is linear and the number of nodes is $10, 20, \ldots, 50$. It

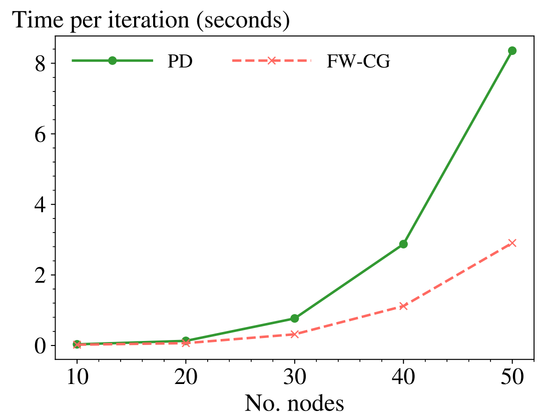


Figure 3.9: The plot shows average time per iteration (seconds) used by PD and FW-CG when demand function is linear and the number of nodes is $10, 20, \ldots, 50$.

can be seen that the running time of PD increases exponentially as the number of nodes increases, while the running time of FW-CG seems to increase linearly as the number of nodes increases. One reason why we observe larger running time in PD is because linear programming subproblems of PD require more time to solve than those of FW-CG due to the number of decision variables generated from column generation method (See Figure 3.10).

Piece-wise Linear Demand Model.

Recall that when we assume demand function to be piece-wise linear, we can show that the revenue function is concave but non-smooth. Since the Frank-Wolfe algorithm with column generation (FW-CG) requires smooth objective function, only the Primal-Dual algorithm (PD) can be applied to solve this problem.

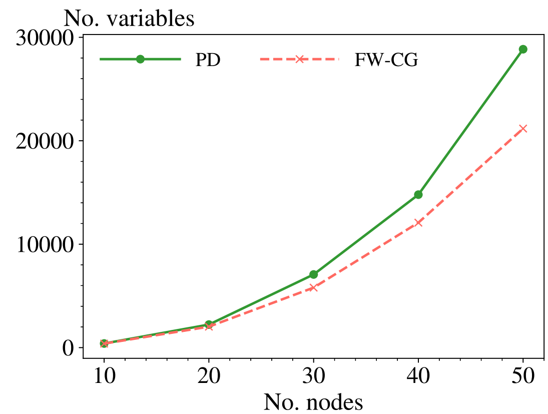


Figure 3.10: The plot shows the number of decision variables involved in linear programming subproblem of PD and FW-CG when demand function is linear and the number of nodes is $10, 20, \ldots, 50$.

We run the Primal-Dual algorithm (PD) on the simulated network for 10000 iterations. The convergence results can be found in Figure 3.11. It can be seen that the slope of log loss and log optimal demand gap in Figure 3.11 scale linearly with log(iteration) which are consistent with Theorem 3 and Theorem 4. When we examine the run time, PD takes 92.95 seconds per 10000 iterations.

3.6 Final Remarks

We have studied an integrated pricing and routing problem on a service network. The problem can be formulated as a convex optimization problem, although the size of this optimization problem prohibits us from solving it directly. We have proposed two algorithms to solve the problem. First, we modified the classical Frank-Wolfe algorithm with column generation (FW-CG). When the objective function is smooth, we show that the rate of convergence in terms of the objective function is $O(1/T)$. Second, we propose a primal-dual algorithm (PD), which allows a non-smooth objective function (e.g., a piece-wise linear pricing function). We show that when the objective is strongly concave, the rate of convergence of PD in terms of average regret is $O(\log T/T)$. Numerical experiments demonstrate the benefit of joint pricing and routing; it can increase profit by more than 10%.

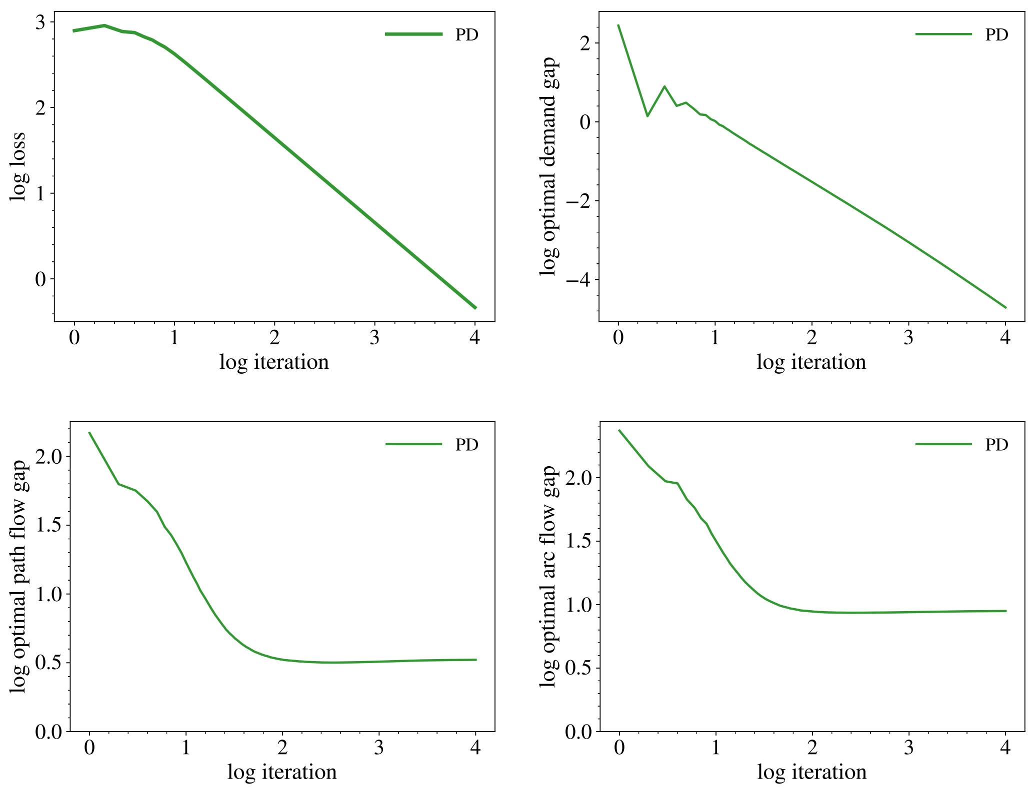


Figure 3.11: The plots compare log loss (upper left), log optimal demand gap (upper right), log optimal path flow gap (lower left) and log optimal arc flows gap (lower right) of PD against log iteration (x-axis) when demand function is piece-wise linear.

There are many possible directions for future research. We mention only two here. We have assumed that the service capacity in the service network is known. A natural extension is to let the optimization decide whether it is beneficial to acquire additional capacity on some of the links in the network. This can be done by allowing demand on an arc to exceed capacity at a cost. Another natural extension is to consider convex arc cost, e.g., per unit cost is an increasing step function as a result of resource scarcity. In this case, we can apply PD to solve the problem; however, the sub-problem becomes a nonlinear MCMCF problem where the objective function is piece-wise linear.

CHAPTER 4

DYNAMIC PRICING AND MATCHING FOR TWO-SIDED QUEUE

4.1 Introduction

Most queueing models consider a fixed set of servers with sequentially arriving customers. In this chapter, we consider a two-sided queueing system where servers also arrive sequentially and then wait to be matched with customers. Several applications of online marketplaces and gig economy platforms can be modeled as two-sided queues—for example, Uber and Lyft where passengers are matched with drivers, Grubhub and DoorDash where customer orders are matched with meal delivery couriers, and crowdsourced workforce platforms such as TaskRabbit where tasks are matched with contributors. Most of these platforms use both dynamic pricing and dynamic matching as levers to control market profitability and delay efficiency.

Motivated by these applications, we consider a canonical model of two-sided queues with multiple types of servers and multiple types of customers. Each customer type is compatible with a certain subset of server types. For example, in the case of ride-hailing marketplaces, the types of servers (drivers) and customers are determined by the proximity of their current locations, as well as other factors such as numbers of seats needed and vehicle capacities. Our model assumes a fairly general setting with arbitrary numbers of customer and server types, with their compatibility modeled by a bipartite graph.

At each point in time, the system operator posts a price for each customer and server type. Then, customers and servers who are willing to accept the quoted prices (after they factor in expected waiting costs) will enter the system. Those who entered will wait in queues separated by their types until they are matched to a compatible counterpart type. Once a customer-server pair is matched, the pair will leave the queueing system immediately in order to complete the service. The system operator earns a profit that is equal to the difference between the price charged to the customer and the price quoted to the server.

We formulate the above system as a Markov decision process (MDP) in the infinite time horizon. The operator can vary the prices for different customer and server types, as well as decide when to match and which customer-server pair to match. The objective is to maximize the long-run average profit obtained by the system operator.

There are several technical challenges to analyze this stochastic system. The first challenge is the *curse of dimensionality* in solving and analyzing the MDP. As the number of customer or server types increases, the dimension of the state space increases exponentially (even when the buffer size of each queue is bounded). It is hence intractable to solve the exact MDP for large-scale systems with multiple types. In this chapter, we propose several approximate policies to obtain near optimal solutions for the MDP.

The second challenge is that the stochastic behavior of the two-sided queueing system is complicated by the interplay between pricing and matching decisions. Our proposed policies use dynamic pricing to ensure the stability of the two-sided queue system, so the arrival rates of customers and servers vary with the queue lengths. As the queue lengths change, the matching decisions among different types are adjusted dynamically, which in turn affects the system state and pricing decisions. As a result, the queue lengths of different types are intricately correlated. The system cannot be decomposed into a set of simple queue and the pricing and matching decisions cannot be decoupled and analyzed separately. To solve this challenge, we use the Lyapunov drift method to analyze the stochastic system as a whole in order to bound the total queue length.

4.1.1 Summary of Results

We first present a fluid model for the two-sided queueing system (Section 4.3) and show that the profit obtained by the fluid model is an upper bound on the achievable profit under any policy. Based on the fluid model, we propose several approximate policies.

We analyze the proposed policies in a *large-scale* setting in which all the arrival rates are scaled by a factor $\eta \in \mathbb{Z}_+$. Under this scaling regime, any policy within $o(\eta)$ of the optimal expected profit is asymptotically optimal. We first consider a fluid pricing policy combined with the max-weight matching algorithm (Tassiulas and Ephremides, 1992). We show that the profit loss of this policy from the fluid solution is $O(\sqrt{\eta})$ (Section 4.3). We then propose a generalization of the fluid pricing policy that uses two prices for each queue type (Kim and Randhawa, 2017). For this two-price policy combined with max-weight matching policy, we show that the profit loss from the fluid profit is reduced to $O(\eta^{1/3})$ (Section 4.4). Furthermore, we prove that for a broad class of pricing policies, using any matching policy will result in a profit loss lower bounded by $\Omega(\eta^{1/3})$ (Section 4.5).

Next, we compare the max-weight matching policy with other matching policies and show its superiority (Section 4.6.2). We show that the max-weight algorithm is *delay optimal* in the limit as $\eta \to \infty$. In particular, under the fluid pricing policy, max-weight matching minimizes the revenue loss due to hitting the queue buffer size threshold. Under the two-price policy, max-weight matching minimizes the expected sum of queue lengths. Using these results, we characterize the profit loss of max-weight matching as a function of the number of customer/server types n: the profit loss scales as $O(\sqrt{n})$ for fluid pricing policy and $O(n^{1/3})$ for two-price policy. In contrast, if one directly applies the solution of the fluid model as a state-independent randomized matching policy, the profit loss can scale as $\Omega(n)$ for either fluid pricing or two-price policies.

Finally, we analyze the structure of the MDP model and propose approximate DP solutions. In some special cases, we are able to show structural properties of the optimal dynamic pricing policy (Appendix C.1.2). In addition, we present an LP-based approximation technique with a constraint generation algorithm to solve the MDP efficiently (Appendix C.1.3).

4.1.2 Literature Review

Dynamic Matching.

Dynamic matching markets have numerous applications such as ride sharing (Banerjee, Freund, and Lykouris, 2017), e-commerce marketplaces like Amazon.com or Ebay, kidney exchange (Roth, Sönmez, and Ünver, 2007; Anderson, Ashlagi, Gamarnik, and Kanoria, 2017), and payment processing networks (Sivaraman, Venkatakrishnan, Ruan, Negi, Yang, Mittal, Fanti, and Alizadeh, 2020). Below, we will discuss previous work involving dynamic matching in the context of two-sided queues.

Caldentey, Kaplan, and Weiss (2009) and Adan and Weiss (2012) considered bipartite matching for two-sided queues on a first-come-first-served basis: each arriving customer is matched to a compatible server who has the earliest arrival time and has not been matched. Under this matching rule, they analyzed steady-state matching rates between certain customer and server types. Furthermore, they deduced the necessary conditions on the frequency of arrivals for stability of the system and also derived the stationary distribution. Gurvich and Ward (2014) analyzed a general multi-sided queuing system, which includes the two-sided queueing system as a special case. Their objective was to minimize holding cost in a finite horizon. They presented a periodic review matching algorithm and showed asymptotic optimality as arrival rates become large.

Hu and Zhou (2018) studied a two-sided matching system similar to ours. Their goal is to maximize the discounted reward obtained by matching customers and servers in a finite horizon, while accounting for the holding costs. They study conditions such that a priority rule is optimal. In addition, they present a matching algorithm based on fluid approximation and show that it is asymptotically optimal. The main distinction of Hu and Zhou (2018) with our work is that they do not consider dynamic pricing. In addition, while they use fluid approximation to generate static (open-loop) matching decisions, we use max-weight algorithm to generate (closed-loop) matching decisions that are adaptive to queue lengths.

Dynamic matching problems were also studied in the context of kidney exchanges albeit in a non-two-sided setting in Anderson, Ashlagi, Gamarnik, and Kanoria (2017) and Akbarpour, Li, and Gharan (2020). Due to legal requirements, pricing is usually not allowed in kidney exchanges. These papers study the value of "batching", i.e., holding compatible matching pairs in hope that better matching will arrive in future. However, both papers find that batching in general does not provide significant benefit.

Dynamic Pricing for Queues.

First we discuss the literature involving dynamic pricing in the context of single-sided queues and then review those involving two-sided queues.

Low (1974a) is one of the earlier works studying dynamic pricing in a single sided queue. The paper considered price dependent customer arrivals with a finite buffer; the rewards include the payment by customers and holding costs incurred by the operator. Monotonicity of the optimal pricing policy is showed. It was later extended to infinite buffer capacity in Low (1974b). Chen and Frank (2001) considered a queuing model with customers who are sensitive to both waiting time and price. They presented structural properties on optimal pricing decisions and monotonicity of optimal bias function. In the context of network services like call centers, Paschalidis and Tsitsiklis (2000) considered a system with finite total resource. They consider different types of price dependent customers arrivals which requests for a fraction of the resource. The objective is to find a pricing policy to maximize revenue. They show multiple structural properties like concavity of value function and monotonicity of optimal policy.

Kim and Randhawa (2017) considers a single server queuing system and studies the benefit of dynamic pricing over static pricing. They assume that the customers are delay sensitive and consider a revenue maximization objective. They present a static pricing policy and a two-price policy, and also provide the rate of convergence of these policies. Our two-price policy considered in Section 4.4 is motivated by the results from Kim and

Randhawa (2017). The method of Kim and Randhawa (2017) involves applying the Taylor series expansion to the revenue function and then bounding the expected steady-state queue length.

The main distinction of Kim and Randhawa (2017) with our work is that they consider a single server queue, whereas we consider a network of two-sided queues with matching decisions. It is non-trivial to generalize the method presented in Kim and Randhawa (2017) to a two-sided queueing network, as an exact analysis of the steady-state distribution is intractable due to the complex interaction among different queues. In addition, unlike the single server setting in Kim and Randhawa (2017), matching decisions play a critical role in our model and cannot be decoupled from the pricing decisions. Aside from establishing asymptotic rates with large arrival rates, we also complement the result in Kim and Randhawa (2017) by showing the advantage of two-price pricing (when combined with appropriate matching policies) for large network sizes.

The joint problem of dynamic pricing and matching was also studied by Özkan and Ward (2020) under the objective of maximizing the number of successful matches. They proposed an asymptotically optimal pricing and matching policy with large arrival rates. The differences with our work are that they proposed static policies based on the fluid model and analyzed the system for a finite time horizon.

A two-sided queueing model with both customer and server arrivals is studied by Nguyen and Stolyar (2018). They consider a setting where the arrival rate of the servers can be controlled. However, the focus in Nguyen and Stolyar (2018) was to establish system stability and process level convergence, while the objective in our model is to maximize profit.

Several recent papers have studied dynamic pricing in the context of ride hailing systems (Yan, Zhu, Korolko, and Woodard, 2019; Besbes, Castro, and Lobel, 2018; Hu, Hu, and Zhu, 2019). Banerjee, Freund, and Lykouris (2017) and Banerjee, Kanoria, and Qian (2018) studied a closed queuing network, where the number of cars in the system is a constant and the customers abandon the system if they are not matched immediately. Banerjee, Freund, and Lykouris (2017) considered a state-independent pricing policy and prove the approximation ratio with respect to optimal pricing policy. (Banerjee, Johari, and Riquelme, 2016) proposed a state-dependent pricing policy and argue that the benefit of dynamic pricing is in the robustness of the performance of the system.

In sum, most of the previous work on dynamic matching is either in the context of single-sided queues or not coupled with revenue optimization. Of the few that consider both of these, the matching policy considered is an open-loop policy. On the other hand, we consider all of these aspects and show the asymptotic optimality under closed-loop matching policies.

Max-Weight Algorithm.

In this work, we apply a max-weight matching algorithm to two-sided queuing systems. This algorithm was first proposed by Tassiulas and Ephremides (1992) in the context of communication networks. After that, the max-weight algorithm and the backpressure algorithm, which is a generalization of the max-weight algorithm, are studied intensively in the literature. The book by Srikant and Ying (2014) provides an excellent summary. The performance of the max-weight algorithm in the context of a switch operating in heavy traffic has been studied by Maguluri and Srikant (2016). The backpressure algorithm was used in the context of online ad matching in Tan and Srikant (2012) and in the context of ride hailing in Kanoria and Qian (2019).

Heavy traffic analysis of the max-weight algorithm in the context of single-sided queue has a long line of literature. One analysis approach is based on fluid limits, diffusion limits and reflected Brownian motion (RBM) (Harrison, 2013). In this approach, the queueing process is studied under an appropriate scaling and the corresponding limiting fluid or diffusion process is shown to converge to a lower dimensional RBM. This phenomenon is called state space collapse (SSC). If the RBM is single dimensional, then it is called complete resource pooling (CRP). Examples on this line of work to study SSC under the max-weight algorithm in the context of single-sided queues are Williams (1998), Stolyar et al. (2004), and Gamarnik, Zeevi, et al. (2006). In this chapter, we employ another approach based on the Lyapunov drift method developed by Eryilmaz and Srikant (2012) and later used by Maguluri and Srikant (2015) for switch systems. We generalize the Lyapunov function for two-sided queues and analyzed the max-weight algorithm under the CRP condition similar to that in Gurvich and Whitt (2009) and Shi, Wei, and Zhong (2019).

4.1.3 Notation

Throughout the chapter, vectors are denoted by boldface letters. Functions applied on vectors are defined entrywise; e.g., $F(\boldsymbol{\lambda})$ is defined to be $(F(\lambda_1),\dots,F(\lambda_m))$. For any two vectors $\mathbf{a}\in\mathbb{R}^n$ and $\mathbf{b}\in\mathbb{R}^m$, we denote the concatenated vector of dimension $n+m$ by $(\mathbf{a},\mathbf{b})$. We denote the n-dimensional vector with all 1's by $\mathbf{1}_n$, and the n-dimensional vector with all 0's by $\mathbf{0}_n$; we omit the subscript n if the sizes of these vectors are clear from the context. If $\mathbf{x}$ and $\mathbf{y}$ are of the same dimension, we use $\langle\mathbf{x},\mathbf{y}\rangle$ to denote the inner product, and $\mathbf{x}\circ\mathbf{y}$ to denote the Hadamard product (i.e., entrywise product). Any inequality $\mathbf{x}\le\mathbf{y}$ is also defined entrywise. We use the superscript "s" to denote variables related to servers and the superscript "c" for variables related to customers. We use $\mathbf{e}_j^{(c)}$ and $\mathbf{e}_i^{(s)}$ to represent unit vectors with a 1 for type j customer and type i server, respectively, and all 0's otherwise.

4.2 Model

We represent the types of customers and servers by a bipartite graph $G(N\cup M,E)$, where N is the set of server types with $|N|=n$, M is the set of customers type with $|M|=m$, and E is the set of edges representing customer and server types that are compatible with each other (see Figure 4.1). A pair $(i,j)\in E$ if and only if a type j customer can be served by a type i server. Each node in the bipartite graph is a queue of customers or servers waiting to be matched with any one of the compatible counterparts. Our convention is to refer to incoming customers as *demand* and incoming servers as *supply*.

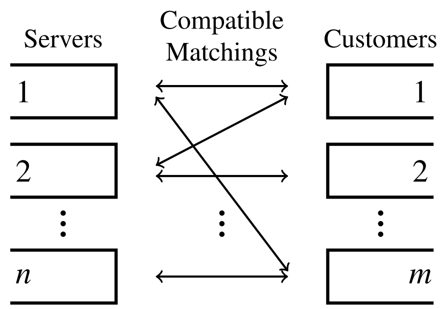


Figure 4.1: Bipartite graph representation for two-sided queues.

At each point in time, the system operator posts a price for each customer and server type. Customers willing to pay the quoted prices, as well as servers who are willing to provide their service at the posted prices (i.e., wages), are admitted to the system. Thus, the system operator can vary the prices to control the arrival rates of customers and servers. Customers and servers then wait in queues until they are matched. The first-come-first-serve (FCFS) discipline is employed for each queue separately, but it may *not* hold among different types of customers and servers. Once a customer is matched with a compatible server, we assume that they depart from the system instantaneously to complete the service process. The system operator's objective is to find a joint pricing and matching policy under which the system is stable (positive recurrent) and the long-run average profit is maximized.

We assume that customers and servers arrive according to nonhomogeneous Poisson processes. For each server type $i \in N$, there exists a supply curve $\mu_i : \mathbb{R}_+ \to \mathbb{R}_+$, such that if the system operator sets a price $p_i^{(s)}$ and the expected waiting time is $w_i^{(s)}$, the resulting arrival rate is $\mu_i\left(p_i^{(s)} - s_i^{(s)} w_i^{(s)}\right)$, where the constant $s_i^{(s)}$ is the unit waiting cost of server type i. Similarly, for each customer type $j \in M$, there exists a demand curve $\lambda_j : \mathbb{R}_+ \to \mathbb{R}_+$, such that if the system operator sets a price $p_j^{(c)}$ and the expected waiting time is $w_j^{(c)}$, the resulting arrival rate is $\lambda_j\left(p_j^{(c)} + s_j^{(c)} w_j^{(c)}\right)$, where $s_j^{(c)}$ is the unit waiting cost of customer type j. We make the following assumption on the supply and demand curves.

Assumption 1. *The supply curves* $\mu_i : \mathbb{R}_+ \to \mathbb{R}_+$ $(\forall i \in N)$ *are strictly increasing and twice* *continuously differentiable. The demand curves $\lambda_j : \mathbb{R}_+ \to \mathbb{R}_+$ $(\forall j \in M)$ are strictly decreasing and twice continuously differentiable.*

Since λ_j and μ_i are strictly monotone, their inverse functions exist, and we denote them by $F_j : \mathbb{R}_+ \to \mathbb{R}_+$ $(\forall j \in M)$ and $G_i : \mathbb{R}_+ \to \mathbb{R}_+$ $(\forall i \in N)$, respectively. In addition, we define the revenue and cost functions as $r_j^{(c)}(\lambda_j) \triangleq \lambda_j F_j(\lambda_j)$ for all $j \in M$ and $r_i^{(s)}(\mu_i) \triangleq \mu_i G_i(\mu_i)$ for all $i \in N$. We make the following assumption on the revenue and cost functions.

Assumption 2. *The revenue function $r_j^{(c)}(\lambda_j)$ is concave $(\forall j \in M)$. The cost function $r_i^{(s)}(\mu_i)$ is convex $(\forall i \in N)$.*

The concavity assumption on revenue function follows from the economic law of diminishing marginal return: as the system operator increases the customer arrival rate λ_j, the marginal revenue $dr_j^{(c)}(\lambda_j)/d\lambda_j$ decreases, which implies that the revenue function $r_j^{(c)}(\lambda_j)$ is concave. This assumption is often assumed in the revenue management literature (Gallego and Van Ryzin, 1994; Kim and Randhawa, 2017). We assume that the marginal cost $dr_i^{(s)}(\mu_i)/d\mu_i$ increases with μ_i, since it becomes harder to recruit servers when we try to increase server arrival rate. This implies that the cost function $r_i^{(s)}$ is convex.

For those customers and servers waiting in queues, the system operator uses matching controls to govern the queueing process. At any given time, suppose $q_i^{(s)}$ is the number of type i servers waiting in queue, and $q_j^{(c)}$ is the number of type j customers waiting in queue. We denote the vector of all queue lengths by $\mathbf{q} = (q_j^{(c)}, \forall j \in M,\ q_i^{(s)}, \forall i \in N)$. We denote the number of type i servers to be matched to type j customers by y_{ij}. The set of feasible matching decisions is

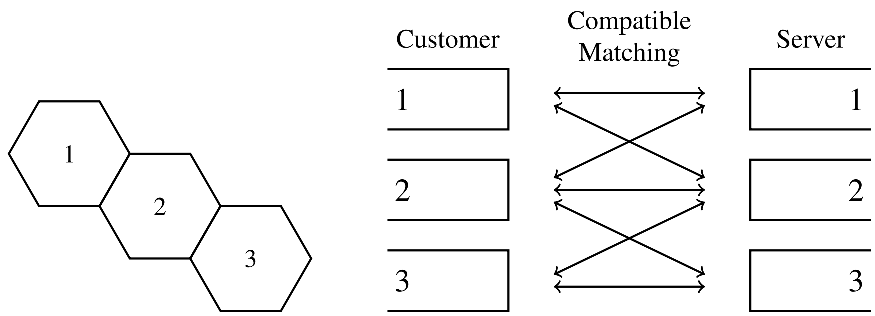
$$Y(\mathbf{q}) \triangleq \Bigg\{ \mathbf{y} \in \mathbb{Z}_+^{nm} \ \Bigg| \ \sum_{i=1}^{n} y_{ij} \le q_j^{(c)}\ (\forall j \in M),\ \sum_{j=1}^{m} y_{ij} \le q_i^{(s)}\ (\forall i \in N),\ y_{ij} = 0\ (\forall (i,j) \notin E) \Bigg\}.$$

We also define the projection of $Y(\mathbf{q})$ to the queue length space as

$$X(\mathbf{q}) \triangleq \left\{ \mathbf{x} \in \mathbb{Z}_{+}^{n+m} \;\middle|\; \exists\, \mathbf{y} \in Y(\mathbf{q}) :\; x_i^{(s)} = \sum_{j=1}^{m} y_{ij}\ (\forall j \in M),\; x_j^{(c)} = \sum_{i=1}^{n} y_{ij}\ (\forall i \in N) \right\}. \tag{4.1}$$

When a pair of customer and server is matched by the system, they both depart from the system. Since a customer is only compatible to a subset of server types, the system operator may have an incentive to hold some customers or servers in queue in order to achieve better matches in future.

Example: Ride Hailing. An application of the two-sided queueing model is in ride hailing systems. In such a system, the customer and server (drivers) types, as well as the matching compatibility graph, are determined by their geographical locations. A simple example with three regions is shown in Figure 4.2. (Here, we ignore issues such as vehicle capacity and number of seats requested by customers, which can be accounted for by creating additional customer and server types.) Based on the price and the waiting time quoted to customers, only a fraction of them who open the app will book a ride, which determines the customer arrival rate. Similarly, based on the price quoted to the drivers, they will choose whether or not to provide service. Thus, the arrival rates of customer and drivers depend on price and wait time and are governed by the demand and supply curve of each region. Once a customer confirms the price and books a ride, the system operator can determine which driver (from what region) should be matched to the customer. If a driver accepts the ride request, then they immediately become unavailable for any other ride requests (departing from the system). After the ride is complete, the car becomes available again, possibly in a different region. A simplifying assumption in our model is that we treat a driver who completes the service and re-enters the system the same as a new arrival.



(a) Geographical locations (b) Bipartite graph

Figure 4.2: A ride hailing system with three regions where we assume that riders can only be matched to cars in their own region or any neighboring regions. The two-sided system generated from the map is shown in Subfigure (b).

4.2.1 Continuous-Time MDP Formulation

We now formulate the system operator's decision problem as a continuous-time Markov decision process (CTMDP) and specify its states, actions, transition rates, and rewards. The system state is represented by the queue lengths of all customer and server types $\mathbf{q} \in \mathbb{Z}_+^{n+m}$. The actions of the CTMDP include both pricing and matching decisions. The matching decision must satisfy $\mathbf{x} \in X(\mathbf{q})$ defined by Eq (4.1). For the pricing decision, in order to leverage Assumption 2, it is more convenient to use arrival rates $(\boldsymbol{\lambda}, \boldsymbol{\mu})$ rather than prices as the control variables. In particular, for customer type $j \in N$, setting the arrival rate to λ_j is equivalent to setting the price to $p_j^{(c)} = F_j(\lambda_j) - s_j^{(c)} w_j^{(c)}(\mathbf{q})$. Similarly, for server type $i \in M$, setting the arrival rate to μ_i is equivalent to setting the price to $p_i^{(s)} = G_i^{(s)}(\mu_i) + s_i^{(s)} w_i^{(s)}(\mathbf{q})$. Thus, the action is a tuple $\mathbf{z} \triangleq (\boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{x}) \in \mathbb{R}^{2(m+n)}$. Given this action, the transition rate from state $\mathbf{q}$ to state $\mathbf{q} + \mathbf{e}_j^{(c)} - \mathbf{x}$ (i.e., having a new arrival of type j customer) is λ_j ($\forall j \in M$), and a reward of $p_j^{(c)}$ is received upon the new arrival. The transition rate from state $\mathbf{q}$ to state $\mathbf{q} + \mathbf{e}_i^{(s)} - \mathbf{x}$ (i.e., having a new arrival of type i server) is μ_i ($\forall i \in N$), and a cost of $p_i^{(s)}$ is paid upon the new arrival. The system operator's objective is to find a pricing and matching policy such that the long run average profit earned by the system operator is maximized. We restrict our attention to policies that make the system stable in the long run, which is defined as follows.

Definition 2. *A joint pricing and matching policy is said to be stable, if the continuous-time Markov chain (CTMC) induced by this policy has a positive recurrent communicating class that contains the state* $\mathbf{q} = \mathbf{0}$.

Remark 1 (Average waiting time.)**.** *It is technically challenging to analyze the exact waiting time* $w_i^{(s)}(\mathbf{q})$ *and* $w_j^{(c)}(\mathbf{q})$, *since the waiting time of one type may depend on the queue lengths of all the types as well as the policy and matching policy used by the system operator. Additionally, in some applications, real-time queue length information may not be visible to all market participants (Zohar, Mandelbaum, and Shimkin, 2002). Therefore, we make a simplifying assumption that the waiting time* perceived *by the customers and servers is the long-run average waiting time. That is, we assume*

$$p_j^{(c)} = F_j(\lambda_j) - s_j^{(c)}\,\mathbb{E}[w_j^{(c)}(\mathbf{q})] \ \ \forall j \in M, \quad p_i^{(s)} = G_i^{(s)}(\mu_i) + s_i^{(s)}\,\mathbb{E}[w_i^{(s)}(\mathbf{q})] \ \ \forall i \in N.$$

The scheme of announcing the long-run average waiting time to (impatient) customers is commonly assumed in the literature (Zohar, Mandelbaum, and Shimkin, 2002; Armony, Shimkin, and Whitt, 2009). Additionally, in the large scale setting that will be considered in the following sections, approximating real-time estimated waiting time with the long-run average waiting time will only result in a negligible error term of a higher order (see Kim and Randhawa (2017), Section 6.1 for a similar argument).

Equivalence to Holding Cost Models.

The above model assumes that customers and servers are sensitive to both prices and waiting costs when they decide to enter the queueing system. We now consider an alternative model, where customers and servers only react to prices, while the system operator pays *additional* holding costs for market participants waiting in queues. In particular, in this alternative model, the states, actions, and transition rates remain the same. Given a state $\mathbf{q}$

and an action $\mathbf{z} = (\boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{x})$, the reward function is defined as

$$\mathscr{R}(\mathbf{q}, \mathbf{z}) \triangleq \sum_{j=1}^{m} \lambda_j F_j(\lambda_j) - \sum_{i=1}^{n} \mu_i G_i(\mu_i) - \sum_{j=1}^{m} s_j^{(c)} q_j^{(c)} - \sum_{i=1}^{n} s_i^{(s)} q_i^{(s)}, \tag{4.2}$$

where $s_j^{(c)}$ and $s_i^{(s)}$ are the customers' and servers' impatience parameters introduced in the original model. The following result shows that the two modelling approaches are indeed equivalent.

Proposition 4. *For any given control policy, the delay-sensitive model and the holding cost model have the same long-run average profit.*

The proof of Proposition 4 follows an application of Little' Law and can be found in Appendix C.1.1. The advantage of considering the holding cost model is that the reward function $\mathscr{R}(\mathbf{q}, \mathbf{z})$ does not explicitly depend on the waiting time. Hence, we use the holding cost model in the rest of the paper.

4.2.2 Discrete-Time MDP Formulation by Uniformization

Instead of analyzing the CTMDP directly, we use the well-known uniformization technique (e.g. Puterman, 1994, Chap. 11) to obtain an equivalent discrete-time Markov decision process (DTMDP), which will simplify our analysis. The uniformized process works as follows. We first choose a uniformization parameter c defined below.

Definition 3. *Suppose there exists constants $\boldsymbol{\lambda}_{\max} \in \mathbb{R}_+^m$ and $\boldsymbol{\mu}_{\max} \in \mathbb{R}_+^n$ such that for any price vector $\mathbf{p} \geq 0$ we have, $\boldsymbol{\lambda}(\mathbf{p}) \leq \boldsymbol{\lambda}_{\max}$, and $\boldsymbol{\mu}(\mathbf{p}) \leq \boldsymbol{\mu}_{\max}$. Let c be any constant such that $c \geq \langle \mathbf{1}_m, \boldsymbol{\lambda}_{\max} \rangle + \langle \mathbf{1}_n, \boldsymbol{\mu}_{\max} \rangle$.*

The uniformized DTMDP is endowed with the same state $\mathbf{q}$ and action $\mathbf{z} = (\boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{x})$ as the CTMDP. Let $Z(\mathbf{q}) = [0, \boldsymbol{\lambda}_{\max}] \cup [0, \boldsymbol{\mu}_{\max}] \cup X(\mathbf{q})$ be the set of feasible actions for queue length $\mathbf{q} \in \mathbb{R}_+^{m+n}$. In the uniformized DTMDP, there is at most one customer arrival or one server arrival in each period: the state transitions from $\mathbf{q}$ to $\mathbf{q}+\mathbf{e}_j^{(c)}-\mathbf{x}$ with probability λ_j/c ($\forall j \in M$); it transitions from $\mathbf{q}$ to $\mathbf{q}+\mathbf{e}_i^{(s)}-\mathbf{x}$ with probability μ_j/c ($\forall i \in N$); otherwise, no arrival happens in this period, and the state remains at $\mathbf{q}$ with probability $1-(\langle \mathbf{1}_m, \boldsymbol{\lambda}\rangle + \langle \mathbf{1}_n, \boldsymbol{\mu}\rangle)/c$. The expected reward in one period is given by $\mathscr{R}(\mathbf{q},\mathbf{z})/c$. Let $\mathbf{q}'$ be the state in the next period. The Bellman equation of the DTMDP is

$$h(\mathbf{q})+\frac{\gamma}{c} \;=\; \max_{\mathbf{z}\in Z(\mathbf{q})}\left\{\frac{\mathscr{R}(\mathbf{q},\mathbf{z})}{c}+\mathbb{E}[h(\mathbf{q}') \mid \mathbf{q},\mathbf{z}]\right\}, \quad \forall \mathbf{q}\in \mathbb{Z}_+^{n+m}, \tag{4.3}$$

where

$$\begin{aligned}\mathbb{E}[h(\mathbf{q}') \mid \mathbf{q},\mathbf{z}] \;=\; & \sum_{j=1}^{m}\frac{\lambda_j}{c}h(\mathbf{q}+\mathbf{e}_j^{(c)}-\mathbf{x})+\sum_{i=1}^{n}\frac{\mu_i}{c}h(\mathbf{q}+\mathbf{e}_i^{(s)}-\mathbf{x}) \\ & +(1-\sum_{j=1}^{m}\frac{\lambda_j}{c}-\sum_{i=1}^{n}\frac{\mu_i}{c})h(\mathbf{q}). \end{aligned} \tag{4.4}$$

In the above equation, the solution γ is the optimal long-run average profit, and $h(\mathbf{q})$ is the bias function associated with state $\mathbf{q}$ ($\forall \mathbf{q}\geq 0$). (Note that the optimal solution of the uniformized DTMDP satisfies that the above Bellman equation because we require the optimal policy to be stable, see Definition 2.) The term $\mathbb{E}[h(\mathbf{q}') \mid \mathbf{q},\mathbf{z}]$ is the expectation of the bias function h after one transition in the uniformized process. The expectation is taken with respect to the one-period transition probabilities conditional on the state $\mathbf{q}$ and the action $\mathbf{z}$.

In Appendix C.1, we present additional analysis of the uniformized DTMDP. We show the monotonicity structure of the optimal pricing policy in the single-link queueing system (i.e., $m=n=1$). Unfortunately, as the number of customer and server types becomes large, solving the DTMDP becomes intractable due to the curse of dimensionality. We propose two approximation methods to obtain near optimal solutions to the DTMDP. The first method is based on fluid approximation. The remainder of the chapter primarily focuses on this approach. The second method uses value function approximation. We defer details of this method to Appendix C.1, because the remaining parts of the chapter do not rely on it.

4.2.3 Max-Weight Matching Policy

In the following sections, we will extensively use the *max-weight* matching policy, so we provide its definition here. Suppose the system has state $\mathbf{q}$ and the set of feasible matches is $X(\mathbf{q})$ (see Eq (4.1)). The policy chooses the matching decision $\mathbf{x}$ to be the solution of

$$\arg\max_{\mathbf{x}\in X(\mathbf{q})} \{\langle \mathbf{q}, \mathbf{x}\rangle\}. \tag{4.5}$$

In other words, under the max-weight policy, when there is either a customer or a server arrival, a match will be made if any of the compatible types has a nonempty queue, and we will always match the arriving customer/server to the compatible type with the most number of customers/servers waiting in queue. Otherwise, if all the compatible counterparts' queues are empty, then the arrival is inserted into the queue of its own type.

The max-weight matching policy, originally proposed by Tassiulas and Ephremides (1992), is extensively studied in the queueing literature. This literature is reviewed in Section 4.1.2. Apart from the queueing literature, in the our model specifically, there is also an alternative way to motivate the max-weight matching policy through quadratic value function approximation of the MDP. Suppose the bias function in Eq (4.4) is approximated by $h(\mathbf{q}) \approx \langle 1, \mathbf{q}^2\rangle$, then the *optimal* matching policy of the DTMDP will be very close to a max-weight policy define in Eq (4.5). Appendix C.1.3 contains a detailed discussion of the value function approximation method.

4.3 Asymptotic Optimality of the Fluid Pricing Policy

In this section, we consider a fluid approximation of the queueing system where random arrivals are replaced by deterministic arrival processes. Based on the fluid model, we pro-

Algorithm 9 Max-Weight Matching Policy

input: current queue length $\mathbf{q}(k)$, new arrival $\mathbf{a}(k)$ # k is a decision epoch
initialization: $\mathbf{y}(k) = \mathbf{0}$
for $i \in N$ **do**
  if $a_i^{(s)}(k) = 1$ **and** $\max_{j:(i,j)\in E} q_j^{(c)} > 0$ **then**
    choose $j^* \in \arg\max_{j:(i,j)\in E} q_j^{(c)}$ (breaking ties arbitrarily)
    set $y_{ij^*}(k) = 1$
for $j \in M$ **do**
  if $a_j^{(c)}(k) = 1$ **and** $\max_{i:(i,j)\in E} q_i^{(s)} > 0$ **then**
    let $i^* \in \arg\max_{i:(i,j)\in E} q_i^{(s)}$ (breaking ties arbitrarily)
    set $y_{i^*j}(k) = 1$
output: matching decision $\mathbf{y}(k)$

pose a static pricing and max-weight matching policy and show that it is asymptotically optimal.

4.3.1 Fluid Model

We consider a deterministic optimization problem to maximize the long-run average profit. Suppose customers arrive with constant rates $\boldsymbol{\lambda}$ and servers arrive with constant rates $\boldsymbol{\mu}$. Let χ_{ij} be the average rate of type i server matched to the type j customer for all $(i,j) \in E$. The fluid model is defined as

$$\gamma^* = \max_{\boldsymbol{\lambda},\boldsymbol{\mu},\boldsymbol{\chi}} \quad \langle F(\boldsymbol{\lambda}), \boldsymbol{\lambda}\rangle - \langle G(\boldsymbol{\mu}), \boldsymbol{\mu}\rangle \tag{4.6a}$$

$$\text{subject to} \quad \lambda_j = \sum_{i=1}^{n} \chi_{ij}, \quad \forall j \in M, \tag{4.6b}$$

$$\mu_i = \sum_{j=1}^{m} \chi_{ij}, \quad \forall i \in N, \tag{4.6c}$$

$$\chi_{ij} = 0, \;\; \forall (i,j) \notin E, \;\; \chi_{ij} \geq 0, \;\; \forall (i,j) \in E. \tag{4.6d}$$

We denote an optimal solution to the above fluid problem by $(\boldsymbol{\lambda}^*, \boldsymbol{\mu}^*, \boldsymbol{\chi}^*)$.

To interpret the fluid model above, note that Eqs (4.6b) and (4.6c) are the balance equations for the number of customers and servers matched. Eq (4.6d) specifies that matching is only allowed among compatible customer-server pairs. Intuitively, it is easy to see that these constraints are necessary, because if the balance equations do not hold, then some customer or server types will keep accumulating over time. Thus, the optimization program Eq (4.6) serves as an *upper bound* on the achievable profit under any pricing and matching policy that makes the system stable. This is formally shown in the following proposition. The proof can be found in Appendix C.2.1.

Proposition 5. *The optimal value of the fluid problem Eq* (4.6) *is an upper bound on the long run expected profit rate under any policy that makes the system stable.*

4.3.2 Fluid Pricing Policy

Using the fluid model, we study the two-sided queueing system in a large-scale regime where the arrival rates of all customer and server types are simultaneously scaled by a factor of $\eta \in \mathbb{N}$.

Definition 4 (Large-Scale Regime)**.** *Consider a family of two-sided queueing systems associated with the same bipartite graph $G(N \cup M, E)$ parametrized by $\eta \in \mathbb{N}$. For the η^{th} system, the demand and supply curves satisfy $F^{\eta}(\eta\boldsymbol{\lambda}) = F(\boldsymbol{\lambda})$ for all $\mathbf{0}_m \leq \boldsymbol{\lambda} \leq \boldsymbol{\lambda}_{\max}$ and $G^{\eta}(\eta\boldsymbol{\mu}) = G(\boldsymbol{\mu})$ for all $\mathbf{0}_n \leq \boldsymbol{\mu} \leq \boldsymbol{\mu}_{\max}$.*

Definition 5 (Profit Loss)**.** *The profit loss (denoted by L^{η}) of a policy is the difference between the optimal value of the (scaled) fluid model, denoted by γ_*^{η}, and the long run average profit (including the penalty incurred due to waiting) under that policy.*

According to Definition 4, it is easily verified that the fluid solution to the η^{th} scaled system is given by $\eta\boldsymbol{\lambda}^*$ and $\eta\boldsymbol{\mu}^*$, where $\boldsymbol{\lambda}^*$ and $\boldsymbol{\mu}^*$ is the optimal solution of the unscaled fluid model Eq (4.6). The optimal value of the η^{th} fluid model is $\gamma_*^{\eta} = \eta\gamma^*$. Therefore, if the profit loss of a policy is sublinear in η, namely $L^{\eta} = o(\eta)$, we say the policy is asymptotically optimal in the large-scale regime.

We use the fluid model to define a fluid pricing policy defined as follows:

$$\lambda_j(\mathbf{q}) = \begin{cases} \lambda_j^* & \text{if } q_j^{(c)} < q_{\max}^{\eta} \\ 0 & \text{otherwise} \end{cases} \quad \forall j \in M, \tag{4.7}$$

$$\mu_i(\mathbf{q}) = \begin{cases} \mu_i^* & \text{if } q_i^{(s)} < q_{\max}^{\eta} \\ 0 & \text{otherwise} \end{cases} \quad \forall i \in N.$$

Here, $q_{\max}^{\eta}$ denotes the maximum queue buffer size; it is a parameter that depends on η, which will be specified later.

The main intuition of the fluid pricing policy is the following. When all queues are below their maximum buffer capacity $\mathbf{q}^{\eta}$, the profit rate of the fluid pricing policy is exactly equal to $\eta\gamma^*$. If any customer queue is full, say, $q_j^{(c)} = q_{\max}^{\eta}$, then all future arrivals to queue j will be rejected until at least one customer waiting in queue j is matched. Thus, a fraction of revenue is lost due to customer rejections. More specially, let γ^{η} be the long run average profit of the fluid pricing policy (excluding waiting costs). Let $\mathbf{I}^{(s)}(q_{\max}^{\eta})$ be a (vector) indicator function representing whether server queues are at the maximum capacity, and let $\mathbf{I}^{(c)}(q_{\max}^{\eta})$ be a (vector) indicator function representing whether customer queues are at the maximum capacity. Then, we have

$$\begin{aligned} L^{\eta} &= \gamma_*^{\eta} - (\gamma^{\eta} - \langle \mathbf{s}, \mathbb{E}[\mathbf{q}] \rangle) \\ &= \eta \left(\langle F(\boldsymbol{\lambda}^*), \boldsymbol{\lambda}^* \rangle - \langle G(\boldsymbol{\mu}^*), \boldsymbol{\mu}^* \rangle \right) - \left\langle F(\boldsymbol{\lambda}^*), \eta \boldsymbol{\lambda}^* \circ (\mathbf{1} - \mathbb{E}[\mathbf{I}^{(c)}(q_{\max}^{\eta})]) \right\rangle \\ &\qquad - \eta \left\langle G(\boldsymbol{\mu}^*), \boldsymbol{\mu}^* \circ (\mathbf{1} - \mathbb{E}[\mathbf{I}^{(s)}(q_{\max}^{\eta})]) \right\rangle + \langle \mathbf{s}, \mathbb{E}[\mathbf{q}] \rangle \\ &= \eta \left(\left\langle F(\boldsymbol{\lambda}^*), (\boldsymbol{\lambda}^* \circ \mathbb{E}[\mathbf{I}^{(c)}(q_{\max}^{\eta})]) \right\rangle - \left\langle G(\boldsymbol{\mu}^*), (\boldsymbol{\mu}^* \circ \mathbb{E}[\mathbf{I}^{(s)}(q_{\max}^{\eta})] \right\rangle \right) + \langle \mathbf{s}, \mathbb{E}[\mathbf{q}] \rangle, \end{aligned} \tag{4.8}$$

where the first equality follows from Definition 5, and the second equality uses the definition of the fluid pricing policy. As a result, Eq (4.8) shows that the profit loss of the fluid pricing policy depends on the parameter $q^{\eta}_{\max}$. If we increase the buffer capacity $q^{\eta}_{\max}$, then the probability of dropping customers/servers will reduce, i.e., $\mathbb{E}[\mathbf{I}(\mathbf{q}^{\eta}_{\max})]$ will decrease. However, increasing the buffer capacity will lead to increasing in the expected queue lengths, which will increase the penalty incurred due to waiting. Thus, we choose buffer capacity to balance the trade-off in order to minimize the overall profit loss. Precisely, we will see that choosing $q^{\eta}_{\max} \sim \sqrt{\eta}$ will result in $\mathbb{E}[\mathbf{I}(\mathbf{q}^{\eta}_{\max})] \sim \eta^{-1/2}$ and $\mathbb{E}[\langle \mathbf{1}_{m+n}, \mathbf{q} \rangle] \sim \sqrt{\eta}$, which attains the optimal profit loss.

Theorem 5. *Suppose a family of two-sided queues is given by the bipartite graph $G(N \cup M, E)$ parameterized by η. The profit loss L^{η} under the fluid pricing policy Eq* (4.7) *and max-weight matching (Algorithm 9) is $O(\sqrt{\eta})$, where $q^{\eta}_{\max} = \gamma\sqrt{\eta}$ for any positive constant γ.*

The proof of Theorem 5 can be found in Appendix C.2.2. In addition, it can be shown that the $O(\sqrt{\eta})$ profit loss rate cannot be improved using any fluid pricing policy. The proof of the proposition below is presented in Appendix C.2.3.

Proposition 6. *For a family of two-sided queues parametrized by η, any fluid pricing policy will have a profit loss L^{η} that is at least $\Omega(\sqrt{\eta})$. The choice of $q^{\eta}_{\max} = \gamma\sqrt{\eta}$ for any positive constant γ provides the optimal profit loss rate $\Theta(\sqrt{\eta})$.*

4.4 Asymptotic Optimality of the Two-Price Policy

A main drawback of the fluid pricing policy is that the prices are not adaptive to the system state. In this section, we consider a policy that uses two different prices for each customer/server type. The proposed two-price policy is built on the two-price policy in Kim and Randhawa (2017) for single server queues. Our contribution lies in a joint analysis of two-price and dynamic matching policies in a two-sided queueing network.

The two-price policy can be viewed as a generalization of the fluid pricing policy. We introduce additional parameters $\boldsymbol{\theta} \in \mathbb{R}^m_+$, $\boldsymbol{\phi} \in \mathbb{R}^n_+$ and $\sigma^{\eta} > 0$, which governs the arrival rates of the customers and servers respectively when the queue length is greater than a certain threshold $\tau^{\eta}_{\max}$. The two-price policy is defined as

$$\lambda_j(\mathbf{q}) = \begin{cases} \eta\lambda_j^* & \textit{if } q_j^{(c)} \le \tau^{\eta}_{\max} \\ \eta\lambda_j^* - \theta_j\sigma^{\eta} & \textit{otherwise} \end{cases} \quad \forall j \in M, \tag{4.9}$$

$$\mu_i(\mathbf{q}) = \begin{cases} \eta\mu_i^* & \textit{if } q_i^{(s)} \le \tau^{\eta}_{\max} \\ \eta\mu_i^* - \phi_i\sigma^{\eta} & \textit{otherwise} \end{cases} \quad \forall i \in N.$$

The policy sets a threshold $\tau^{\eta}_{\max}$ for all customer and server types. It uses the fluid arrival rates when queue lengths are below this threshold, and then reduces the arrival rates by $\theta_j\sigma^{\eta}$ outside this threshold for type j customer. Similarly, the policy reduces the server arrival rates outside the threshold by $\phi_i\sigma^{\eta}$ for type i server. Here, $\tau^{\eta}_{\max}$, σ^{η}, $\boldsymbol{\theta}$ and $\boldsymbol{\phi}$ are parameters that will be specified later. (Our convention is to use superscript η to denote any parameter or quantity that is associated with the η^{th} scaled system.) Intuitively, for any type of customer/server, if we increase σ^{η}, the queue length will have a larger negative drift when it exceeds the threshold $\tau^{\eta}_{\max}$, so the expected queue length $\mathbb{E}[\langle \mathbf{1}_{m+n}, \mathbf{q}\rangle]$ will be smaller. However, if σ^{η} are too large, the arrival rates outside the threshold $\tau^{\eta}_{\max}$ will be far from the optimal fluid arrival rates, which will result in a larger profit loss. Thus, there is a trade-off between the expected queue length and profit loss. For the matching algorithm associated with the two-price policy, here we use the max-weight matching algorithm as defined in Eq (4.5). (Other matching algorithms will be considered in Section 4.6.2.) The following theorem provides a bound on the asymptotic performance of the two-price policy as η tends to infinity.

Theorem 6. *Consider a family of two-sided queues parametrized by η represented by the bipartite graph $G(N \cup M, E)$. The profit loss L^{η} under the two-price policy Eq* (4.9) *and the max-weight matching (Algorithm 9) is $O(\eta^{1/3})$ for any $\tau^{\eta}_{\max} \le \eta^{1/3}$, $\sigma^{\eta} = \eta^{2/3}$ and constants $\boldsymbol{\theta} > \mathbf{0}_m$, $\boldsymbol{\phi} > \mathbf{0}_n$.*

The above theorem shows that the profit loss of the two-price policy is $O(\eta^{1/3})$, which is better than the $O(\sqrt{\eta})$ loss in the fluid pricing policy. The proof of the theorem contains two main steps. The first step is to show that the system is stable under the two-price policy and the expected queue lengths are bounded. We also give an upper bound of the expected queue lengths (Lemma 1). The second step in the proof is to estimate the profit loss L^η (Lemma 2) by applying the KKT conditions of the fluid problem.

Lemma 1. *For a system of two-sided queues operating under the two-price policy and the max-weight matching algorithm parameterized by η, the system is positive recurrent for any $\boldsymbol{\theta} > \mathbf{0}_m$, $\boldsymbol{\phi} > \mathbf{0}_n$, $\sigma^\eta > 0$ and $\tau^\eta_{\max} > 0$. The expected queue lengths are bounded by*

$$\mathbb{E}\left[\left\langle \boldsymbol{\theta}, \mathbf{q}^{(c)} \right\rangle\right] + \mathbb{E}\left[\left\langle \boldsymbol{\phi}, \mathbf{q}^{(s)} \right\rangle\right] \leq \tau^\eta_{\max} \left(\sum_{j=1}^{m} \theta_j \Pr[q_j^{(c)} > \tau^\eta_{\max}] + \sum_{i=1}^{n} \phi_i \Pr[q_i^{(s)} > \tau^\eta_{\max}] \right) + \frac{\eta}{\sigma^\eta} \left(\langle \mathbf{1}_n, \boldsymbol{\mu}^* \rangle + \langle \mathbf{1}_m, \boldsymbol{\lambda}^* \rangle \right).$$

Lemma 2. *For a system of two-sided queues operating under two-price policy and max-weight matching policy, for any $\boldsymbol{\theta} > \mathbf{0}_m$, $\boldsymbol{\phi} > \mathbf{0}_n$ and $\tau^\eta_{\max} > 0$, we have*

$$\sum_{j \in M} \left(F_j'(\lambda_j^*)\lambda_j^* + F_j(\lambda_j^*) \right) \theta_j \Pr[q_j^{(c)} > \tau^\eta_{\max}] - \sum_{i \in N} \left(G_i'(\mu_i^*)\mu_i^* + G_i(\mu_i^*) \right) \phi_i \Pr[q_i^{(s)} > \tau^\eta_{\max}] \leq |E|^2 \max_{i \in N, j \in M} \{\phi_i, \theta_j\} \frac{\sigma^\eta}{\eta}.$$

4.5 Lower Bound

In this section, we will obtain lower bounds on the profit loss under a broad family of policies, and thus establish that the $O(\eta^{1/3})$ rate obtained by the two-price policy in Theorem 6 is optimal. In particular, we consider a family of pricing policies that have the following form:

$$\lambda_j = \eta\lambda_j^* + f_j\left(\frac{\mathbf{q}}{\eta^\alpha}\right)\eta^\beta \quad \forall j \in M, \tag{4.10}$$

$$\mu_i = \eta\mu_i^* + g_i\left(\frac{\mathbf{q}}{\eta^\alpha}\right)\eta^\beta \quad \forall i \in N. \tag{4.11}$$

The motivation for this policy is as follows. The first terms in Eqs (4.10) and (4.11) (i.e., $\eta\lambda_j^*$ and $\eta\mu_i^*$) are static and result from the solution of the fluid model; the second terms account for dynamic adjustments as the queue length changes. We assume the adjustment terms can be further decomposed into two terms: a function that scales the queue length, $f_j(\cdot)$ or $g_i(\cdot)$, and a term that determines the scaling of price adjustments, η^β, for some $1 > \beta > 0$. Moreover, as the arrival rates are scaled up, the queue length will be asymptotically large. Thus, we also scale the queue length in function $f_j(\cdot)$ and $g_i(\cdot)$ for all $i \in N$ and $j \in M$ by η^α for some $1 \geq \alpha \geq 0$.

In addition, we require the pricing policy to satisfy the following conditions.

Condition 1. *(a) There exists $\mathbf{M} \in \mathbb{R}_+^m$ and $\mathbf{N} \in \mathbb{R}_+^n$ such that $|f_j(\mathbf{q}/\eta^\alpha)| \leq M_j$ for all $j \in M$ and $|g_i(\mathbf{q}/\eta^\alpha)| \leq N_i$ for all $i \in N$ for all $\mathbf{q} \in S$ and for all $\eta \geq 1$.*

(b) It holds that $0 < \alpha + \beta \leq 1$.

(c) There exists $K > 0$ and $\delta > 0$ such that for all $j \in M$, if $q_j^{(c)}/\eta^\alpha > K$, then either $f_j(\mathbf{q}/\eta^\alpha) < -\delta$ or there exists $i : (i, j) \in E$ such that $g_i(\mathbf{q}/\eta^\alpha) > \delta$ for all η. Similarly for all $i \in N$, if $q_i^{(s)}/\eta^\alpha > K$, then either $g_i(\mathbf{q}/\eta^\alpha) < -\delta$ or there exists $j : (i, j) \in E$ such that $f_j(\mathbf{q}/\eta^\alpha) > \delta$ for all η.

We now interpret the conditions above.Condition 1(a) requires the functions f and g to be bounded, given appropriately scaling of the queue lengths $\mathbf{q}$ as η increases. Condition 1(b) states that the rate of queue length re-scaling (α) should not exceed the rate of re-scaling pricing adjustment terms ($1-\beta$). This condition is needed so that the price adjustment terms are sufficiently large to make the system stable. (In the special case of single-link system, this assumption is not needed; the extension is presented later in Proposition 7.) Condition 1(c) states that if a queue is too long, we should either decrease the arrival rate of this queue or increase the arrival rates of those matched to this queue.

Aside from the above conditions, the pricing forms in Eqs (4.10-4.11) are fairly general, because the pricing function of any queue can depend on the entire system state vector $(\mathbf{q})$, and we do not make any strong assumptions such as monotonicity, continuity or differentiablity on functions f and g. Finally, note that we do not require any assumption on the form of matching policies.

The two-price policy in Section 4.4 satisfies the above condition with

$$f_j(\mathbf{q}) = -\theta_j \mathbb{1}_{q_j^{(c)} > \tau_{\max}} \ (\forall j \in M), \quad g_i(\mathbf{q}) = -\phi_i \mathbb{1}_{q_i^{(s)} > \tau_{\max}} \ (\forall i \in N), \quad \beta = 2/3. \tag{4.12}$$

Now we present the result on the lower bound.

Theorem 7. *For a two-sided queue defined by a graph $G(N \cup M, E)$ operating under any pricing policy of the form Eq* (4.10) *and* (4.11) *that satisfies Condition 1, if the resulting system is stable, there exists a constant $K(F, G, f, g)$ such that*

$$L^{\eta} \ \geq \ K\eta^{1/3}.$$

The details of the proof are deferred to Appendix C.4.1. We present below an intuitive explanation of the rate in the lower bound.

Remark 2 (Intuitive explanation of $\eta^{1/3}$)**.** *The main reason we obtain $O(\eta^{1/3})$ profit loss is due to the trade-off between the expected queue length and the loss in revenue. Consider a pricing policy that deviates from the fluid optimal pricing policy by $\varepsilon > 0$, that is, for all $\mathbf{q} \in S$, we have $|\lambda_j(\mathbf{q}) - \lambda_j^*| < \varepsilon$ for all $j \in M$ and $|\mu_i(\mathbf{q}) - \mu_i^*| < \varepsilon$ for all $i \in N$. One can show that under such a policy, the expected queue length is of the order $1/\varepsilon$ and revenue loss is of the order $\eta\varepsilon^2$. Specifically, the queue length can be coupled to that of an $M/M/1$*

queue in heavy-traffic with parameter ε, whose mean queue length is known to be of the order $1/\varepsilon$ by the Kingman's bound. The loss in revenue can be estimated by the Taylor series expansion of the revenue function. Since we are operating closed to the optimal price of the fluid model, the first order term vanished, and the dominant term of the second order, viz., $\eta\varepsilon^2$. The co-efficient of this term is shown to be strictly positive by analyzing the tail probabilities. Therefore, we have

$$\mathbb{E}\left[\langle \mathbf{1}_{n+m}, \mathbf{q}\rangle\right] \sim \frac{1}{\varepsilon}, \quad L^\eta \sim \eta\varepsilon^2.$$

To achieve the optimal trade-off between expected queue length and profit loss, we choose $\varepsilon \sim \eta^{-1/3}$, which results in the $\eta^{1/3}$ profit loss in Theorem 7.

We can further relax Condition 1(b) in the special case of single-link systems ($m = n = 1$) operating under any two-price policy. The result is stated below and the proof can be found in Appendix C.4.2.

Proposition 7. *For a family of single-link two-sided queue parametrized by η, any two pricing policy given by Eq* (4.9) *with $\sigma^\eta = \eta^\beta$ for some $\beta < 1$ and $\tau^\eta_{\max} = \eta^\alpha$ for some $\alpha \in \mathbb{R}_+$, will have a profit loss L^η at least $\Omega(\eta^{1/3})$. The choice of $\tau^\eta_{\max} = \eta^{1/3}$ and $\sigma^\eta = \eta^{2/3}$ and any positive constants θ and ϕ provides the optimal profit loss $\Theta(\eta^{1/3})$.*

4.6 Further Analysis on Max-Weight Matching

In this section, we present further insights into the max-weight matching algorithm in two-sided queues. First, we show that under the fluid pricing policy, max-weight matching minimizes the probability of hitting the queue length threshold among all matching policies. Second, under the two-price policy, we show that max-weight matching minimizes the expected sum of queue lengths among all possible matching policies. Third, we compare max-weight matching with a randomized matching policy with probabilities specified by the fluid model and show that max-weight has smaller loss in terms of the number of customer/server types. Together, these results show the superiority of the max-weight policy.

We start by establishing *state space collapse* under max-weight matching. State space collapse means that all the customer queues are almost equal in length and all the server queues are almost equal in length; hence, with high probability, only customers or only servers are waiting in the system. This implies that max-weight ends up matching the maximum possible number of customer-server pairs, as only the excess customers/servers are waiting in the system. To achieve state space collapse in the sense discussed above, we propose a complete resource pooling condition on the graph. Similar conditions have been proposed for single-sided queues Gurvich and Whitt (2009, Assumption 2.4) Shi, Wei, and Zhong (2019, Definition 1).

Condition 2 (Complete Resource Pooling (CRP))**.** *There exists an optimal solution* $(\boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)$ *to the fluid problem Eq* (4.6) *such that for all* $J \subsetneq M$ *and for all* $I \subsetneq N$*, it holds that*

$$\sum_{j \in J} \lambda_j^* < \sum_{i: \exists j \in J, (i,j) \in E} \mu_i^*, \quad \sum_{i \in I} \mu_i^* < \sum_{j: \exists i \in I, (i,j) \in E} \lambda_j^*.$$

It is straightforward to verify that the CRP condition implies the connectedness of the graph $G(N \cup M, E)$. The CRP condition also implies that the optimal solution of the fluid problem is in the interior of the feasible region. The following lemma formalizes this observation.

Lemma 3. *If Condition 2 is satisfied, there exists* $\boldsymbol{\chi}^* \geq \mathbf{0}$ *such that* $\chi_{ij}^* > 0$ *for all* $(i, j) \in E$ *and* $(\boldsymbol{\lambda}^*, \boldsymbol{\mu}^*, \boldsymbol{\chi}^*)$ *is an optimal solution to the fluid problem Eq* (4.6)*.*

(The proofs for this section can be found in Appendix C.5.) The above result is not surprising as it is known in the heavy traffic literature (Eryilmaz and Srikant, 2012; Lange and Maguluri, 2019) that if the arrival rate is approaching a point on the boundary of the capacity region in the interior of a facet, then the system exhibits complete resource pooling. However, the analysis of state space collapse for two-sided queues does not follow immediately from the literature of single-sided queues and is more involved. We propose a Lyapunov function approach and use the drift method to show state space collapse. To simplify the analysis, in this section we restrict to a setting where $m = n$ and there exists a perfect matching in the graph $G(N \cup M, E)$.

Condition 3. *The graph $G(N \cup M, E)$ has a perfect matching. Without loss of generality, we assume that server type i is connected to customer type i for all $i \in [n]$.*

In general if $m \neq n$, and if the above condition is not satisfied, we show in Appendix C.5.5 that the pricing and matching problem under a given general graph can be reformulated as a problem under a new graph where the above condition is satisfied. Thus, the results in the following propositions and the theorem can be applied (with minor modifications as shown in Appendix C.5.5) even when the above condition does not hold.

4.6.1 Delay Optimality

We first show the delay optimality of the max-weight matching algorithm among all possible matching algorithms under the fluid pricing policy.

Proposition 8. *Under the fluid pricing policy with any matching algorithm, we have*

$$q_{\max}^{\eta} \left(\sum_{j=1}^{n} \lambda_j^* \Pr\left[q_j^{(c)} = q_{\max}^{\eta} \right] + \sum_{i=1}^{n} \mu_i^* \Pr\left[q_i^{(s)} = q_{\max}^{\eta} \right] \right) \geq \frac{\langle \mathbf{1}_n, \boldsymbol{\lambda}^* \rangle + \langle \mathbf{1}_n, \boldsymbol{\mu}^* \rangle}{2n + 1/q_{\max}^{\eta}}.$$

Furthermore, under the fluid pricing policy with the max-weight matching algorithm, if $q_{\max}^{\eta} \to \infty$ as $\eta \to \infty$, we have

$$\lim_{\eta \to \infty} q_{\max}^{\eta} \left(\sum_{j=1}^{n} \lambda_j^* \Pr\left[q_j^{(c)} = q_{\max}^{\eta} \right] + \sum_{i=1}^{n} \mu_i^* \Pr\left[q_i^{(s)} = q_{\max}^{\eta} \right] \right) = \frac{\langle \mathbf{1}_n, \boldsymbol{\lambda}^* \rangle + \langle \mathbf{1}_n, \boldsymbol{\mu}^* \rangle}{2n}.$$

The above proposition states that the max-weight algorithm (asymptotically) minimizes the proportion of time spent in the threshold state among all possible matching algorithms, hence minimizing the revenue loss caused by hitting the queue length thresholds.

Similarly, the max-weight matching algorithm is delay optimal under the two-price policy. The following proposition states that the max-weight algorithm (asymptotically) minimizes the expected total queue length under the two-price algorithm among all possible matching algorithms.

Proposition 9. *Under the two pricing policy with $\boldsymbol{\theta} = \boldsymbol{\phi} = \mathbf{1}_n$ and any matching policy, the expected total queue length satisfies*

$$\frac{\sigma^\eta}{\eta}\mathbb{E}[\langle \mathbf{1}_{2n}, \mathbf{q}\rangle] \geq \frac{\langle \mathbf{1}_n, \boldsymbol{\lambda}^*\rangle + \langle \mathbf{1}_n, \boldsymbol{\mu}^*\rangle}{2n}.$$

Furthermore, under the two-price policy with $\boldsymbol{\theta} = \boldsymbol{\phi} = \mathbf{1}_n$ and the max-weight matching policy, if $\lim_{\eta\to\infty} \sigma^\eta/\eta = 0$ *and* $\lim_{\eta\to\infty} \sigma^\eta \tau^\eta_{\max}/\eta = 0$, *we have*

$$\lim_{\eta\to\infty} \frac{\sigma^\eta}{\eta}\mathbb{E}[\langle \mathbf{1}_{2n}, \mathbf{q}\rangle] = \frac{\langle \mathbf{1}_n, \boldsymbol{\lambda}^*\rangle + \langle \mathbf{1}_n, \boldsymbol{\mu}^*\rangle}{2n}.$$

Notice that the queue length bound in Proposition 9 is tighter than the bound in Lemma 1, because former requires the CRP condition (Condition 2) whereas the latter does not require such condition. Together, Propositions 8 and 9 establish the asymptotic delay optimality of the max-weight algorithm.

4.6.2 Max-Weight versus Randomized Matching

In this section, we compare the max-weight policy with a randomized matching policy (defined in Algorithm 10) resulting from the fluid model. The randomized matching algorithm matches an incoming arrival to compatible types at fixed probabilities, which are determined by the fluid solution χ^* (see Eq (4.6)). If some queues are empty, the probabilities are rescaled proportionally to match only nonempty queues. Unlike the max-weight algorithm, the randomized matching algorithm does not use information about the queue lengths (except for the emptiness of the queues).

We analyze the profit losses of these two matching algorithms and its dependence on

Algorithm 10 Randomized Matching (Nonempty Queues First)

input: new arrival $\mathbf{a}(k)$, queue length $\mathbf{q}(k)$, the fluid solution χ^* # k is a decision epoch

initialization: $\mathbf{y}(k) = \mathbf{0}$

for $i \in N$ **do**

  if $a_i^{(s)}(k) = 1$ **then**

    set $y_{ij}(k) = 1$ with probability $\frac{\chi^*_{ij}\mathbb{1}_{\{q_j^{(c)}>0\}}}{\sum_{j'=1}^{m}\chi^*_{ij'}\mathbb{1}_{\{q_{j'}^{(c)}>0\}}}$ for all $j \in M$.

for $j \in M$ **do**

  if $a_j^{(c)}(k) = 1$ **then**

    set $y_{ij}(k) = 1$ with probability $\frac{\chi^*_{ij}\mathbb{1}_{\{q_i^{(s)}>0\}}}{\sum_{i'=1}^{n}\chi^*_{i'j}\mathbb{1}_{\{q_{i'}^{(s)}>0\}}}$ for all $i \in N$.

output: matching decision $\mathbf{y}(k)$

the number of customer/server types n when $\eta \to \infty$. First, we consider the fluid pricing policy. The theorem below shows that even though both max-weight and randomized matching have $O(\eta^{1/2})$ profit loss, max-weight matching is order $n^{1/2}$ better than randomized matching policy.

Theorem 8. *Suppose a family of two-sided queues is given by the bipartite graph $G(N \cup M, E)$ parametrized by η. Under the fluid price policy Eq* (4.7) *and randomized matching policy (Algorithm 10), for $q^\eta_{\max} = \gamma\eta^{1/2}$, we have $L^\eta = O(\eta^{1/2})$. For any $\gamma > 0$, there exists $(\boldsymbol{\lambda}^*, \boldsymbol{\mu}^*, \boldsymbol{\chi}^*)$ satisfying Condition 2 and 3 such that*

$$\liminf_{\eta\to\infty} \frac{L^\eta}{\eta^{1/2}} = \Omega(n).$$

In addition, under the fluid price policy (4.7) *and max-weight matching* (4.5) *for $q^\eta_{\max} = \sqrt{\eta/n}$, we have*

$$\limsup_{\eta\to\infty} \frac{L^\eta}{\eta^{1/2}} \le n^{1/2}\left(\frac{\langle \mathbf{1}_n, \boldsymbol{\lambda}^*\rangle + \langle \mathbf{1}_n, \boldsymbol{\mu}^*\rangle}{2n}\max_{j\in N} F_j(\lambda_j^*) + 2\max_{i\in N, j\in M}\{s_i^{(s)}, s_j^{(c)}\}\right) = O(n^{1/2}).$$

Next, we compare the max-weight and randomized matching algorithms for the two-price pricing policy. The theorem below shows that both algorithms achieve $O(\eta^{1/3})$ profit loss, whereas max-weight is order $n^{2/3}$ better than randomized matching.

Theorem 9. *Suppose a family of two-sided queues is given by the bipartite graph $G(N \cup M, E)$ parametrized by η. Under the two-price policy Eq* (4.9) *and randomized matching policy (Algorithm 10), for $\sigma^\eta = \eta^{2/3}$ and $\tau^\eta_{\max} = \gamma\eta^{1/3}$, we have $L^\eta = O(\eta^{1/3})$. For any choice of $\boldsymbol{\theta} > \mathbf{0}$, $\boldsymbol{\phi} > \mathbf{0}$ and $\gamma > 0$, there exists $(\boldsymbol{\lambda}^*, \boldsymbol{\mu}^*, \boldsymbol{\chi}^*)$ satisfying Condition 2 and 3 such that*

$$\liminf_{\eta\to\infty} \frac{L^\eta}{\eta^{1/3}} = \Omega(n).$$

In addition, under the two-price policy Eq (4.9) *and max-weight matching Eq* (4.5) *with $\boldsymbol{\theta} = \mathbf{1}_n$ and $\boldsymbol{\phi} = \mathbf{1}_n$, $\sigma^\eta = n^{-1/3}\eta^{2/3}$, if $\lim_{\eta\to\infty} \tau^\eta_{\max}/\eta^{1/3} = 0$, we have*

$$\limsup_{\eta\to\infty} \frac{L^\eta}{\eta^{1/3}} \le \left(\sum_{i\in N} \left(\frac{\mu_i^* G_i''(\mu_i^*)}{2} + G_i'(\mu_i^*) \right) - \sum_{j\in N} \left(\frac{\lambda_j^* F_j''(\lambda_j^*)}{2} + F_j'(\lambda_j^*) \right) \right.$$
$$\left. + \frac{\max_{i\in N, j\in M}\{s_i^{(s)}, s_j^{(c)}\}}{2} \left(\langle \mathbf{1}_n, \boldsymbol{\lambda}^*\rangle + \langle \mathbf{1}_n, \boldsymbol{\mu}^*\rangle\right) \right) n^{-2/3} = O(n^{1/3}).$$

4.7 Numerical Experiments

4.7.1 Single-Link Systems

Our first experiment analyzes a single-link system with one server type and one customer type. In this case, the system state of the MDP is represented by a single variable, namely, the difference between the customer queue length and the server queue length (a detailed discussion of this system is included in Appendix C.1.2). We will solve the optimal policy of the MDP and compare it with the fluid pricing policy and two-price policy.

We assume a supply curve given by $p_1 = \lambda^{0.5}$ and a demand curve given by $p_2 = 4\mu^{-0.5}$. With these supply and demand curves, the optimal profit of the fluid model is 3.08 when $\lambda = \mu = 4/3$, $p_1 = 1.15$ and $p_2 = 3.46$. We then calculate the optimal pricing policy of the long-run average cost MDP using relative value iteration. Figure 4.3 shows the optimal pricing policy under three different values of the penalty coefficient (s), as well as the optimal price of the fluid model. The result shows that the optimal customer price is always above the server price, and both prices are increasing with the queue length difference. Intuitively, if the system has more customers, the customer price should be increased to reduce the customer arrival rate and server price should be increased to increase the server arrival rate. This observation verifies Proposition 12 in Appendix C.1.2. As s increases, more weight is given to the waiting cost (or equivalently, customers and servers become more sensitive to delays), so the price increases more steeply as the number of customers and servers waiting in the system increases. Figure 4.4 show the stationary distribution and the mean of queue length for different values of the penalty coefficient (s). As expected, when s increases, the queue length is more concentrated around 0. Furthermore, we simulate

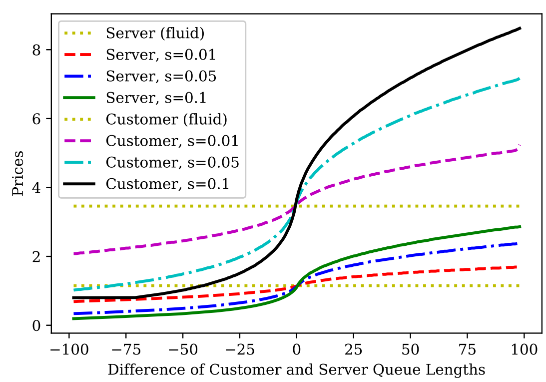


Figure 4.3: Optimal pricing policies under different values of penalty coefficients.

the profit loss under the fluid pricing policy and two-price policy and compare it with the theoretical result presented before and also with the exact solution obtained by solving the MDP. The result is presented in Figure 4.5. The profit loss under the fluid pricing policy has an order of $\sqrt{\eta}$ and that under the two-price policy has an order of $\eta^{1/3}$, verifying Theorem 5 and Theorem 6. Also observe that the profit loss under the two-price policy is not much different from that of the optimal profit loss, demonstrating the effectiveness of a two-price policy.

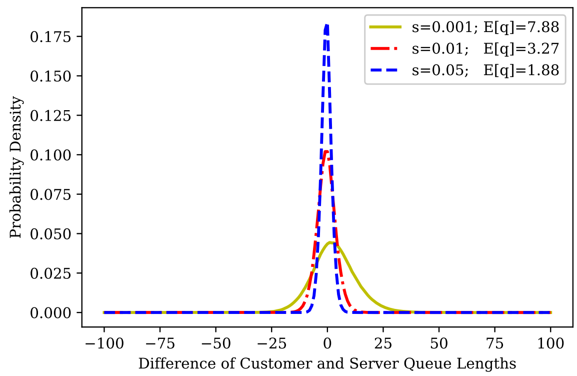


Figure 4.4: Stationary distribution of queue length under different penalty coefficients.

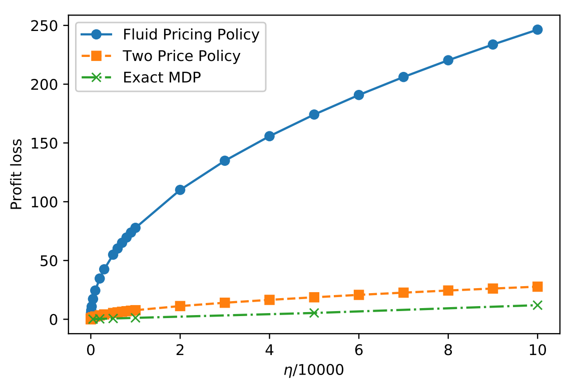


Figure 4.5: Performance of two-price and fluid pricing policy compared to the exact solution.

4.7.2 Systems with Multiple Types

Next, we analyze the general two-sided queues with multiple customer and server types. We examine the profit losses under the following four different algorithms:

1. FP+MW represents the fluid pricing (Eq (4.7)) and max-weight matching (Eq (4.5)) policy;

2. FP+Rand represents the fluid pricing (Eq (4.7)) and randomized matching (Algorithm 10) policy;

3. TP+MW represents the two-price policy (Eq (4.9)) with max-weight matching (Eq (4.5));

4. TP+Rand represents the two-price policy (Eq (4.9)) with randomized matching (Algorithm 10).

In this numerical experiment, we first consider a setting where the number of servers and the number of customers are equal ($m = n$) and CRP condition (Condition 2) is satisfied. We assume the compatibility graph is given by

$$E = \{(i, j) \in [n] \times [n] : j \in \{i+k\} \cup \{(i+k-n)^+\}, k = 0, 1, 2, 3\}.$$

The demand and supply curves are given by

$$F_j(\lambda_j) = 2 - \lambda_j/2, \ \forall j \in [m], \quad \text{and} \quad G_i(\mu_i) = \mu_i/2 \ \forall j \in [n],$$

respectively. We assume the unit holding cost is $s = 1$. The parameter of the fluid pricing policy is set to $q^{\eta}_{\max} = 2\sqrt{\eta/n}$. The parameters of the two-price policy are chosen to be $\tau^{\eta}_{\max} = 0$ and $\sigma^{\eta} = \eta^{2/3} n^{-1/3}$.

We report the profit loss for $\eta \in \{10, 100, 500, 1000, 2000, 5000, 10000\}$ when $m = n = 6$ in Figure 4.6 (left). We find that when η is larger, the profit loss of TP+MW grows the slowest, followed by the profit loss of TP+Rand, FP+MW, and FP+Rand. This result confirms the advantage of the two-price policy over the fluid pricing policy, as well as the advantage of the max-weight matching policy over the randomized matching policy. Figure 4.6 (right) shows the same plot in logarithmic scale. Note that the slope of log-log plot in Figure 4.6 (right) can be interpreted as the order of profit loss with respect to η. The fitted slopes of FP+MW and TP+MW are 0.51 and 0.33 respectively. This is consistent with Theorem 5 and Theorem 6, which state that FP+MW and TP+MW have the orders of profit loss with respect to η of $O(\eta^{1/2})$ and $O(\eta^{1/3})$ respectively. Figure 4.6 shows that FP+MW and FP+Rand yield the same order of profit loss with respect to η of approximately 1/2. Moreover, the two-price policy combined with either max-weight matching or randomized matching yields the same order of profit loss with respect to η of approximately 1/3. That is, choosing max-weight or randomized matching does not affect order of profit loss with respect to η.

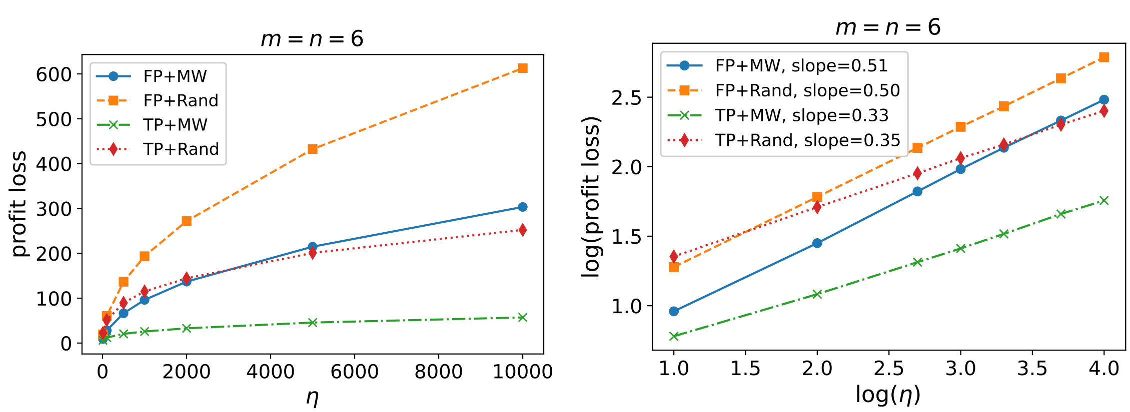


Figure 4.6: Profit losses under FP+MW, FP+Rand, TP+MW and TP+Rand for for different η (left) and its associated log-log plot (right) when $m = n = 6$.

Our next experiment investigates how the profit loss changes with the number of customer and server types. We first consider a setting when the types of two sides are balanced ($m = n$). Figure 4.7 shows the profit loss for $n \in \{4, 6, 8, \dots, 20\}$ when $\eta = 10000$ (a large η is chosen so that the asymptotic trend becomes clear) and its associated plot in logarithmic scale. It can be observed that the profit losses when a pricing policy is combined with the randomized matching policy grow faster than those when a pricing policy is combined with the max-weight matching policy as n increases. In other words, the max-weight matching policy performs better than the randomized matching policy. Figure 4.7 suggests that the orders of profit loss with respect to n of FP+MW and FP+Rand are 0.49 and 1.18 respectively, which are close to those predicted by Theorem 8. Moreover, the orders of profit loss with respect to n of TP+MW and TP+Rand are 0.34 and 1.28, which are close to those predicted by Theorem 9. Clearly, in both cases, the max-weight algorithm performances much better than the randomized matching algorithm for large n.

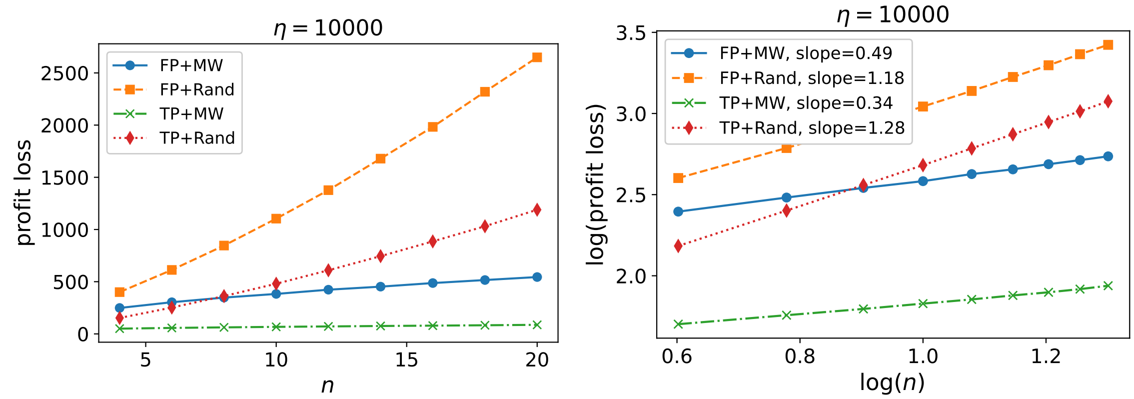


Figure 4.7: Profit losses under FP+MW, FP+Rand, TP+MW and TP+Rand for different n (left) and its associated log-log plot (right) when $\eta = 10000$.

We also consider the setting where the number of server queues and the number of customer queues are not equal. Specifically, we assume that the number of server queues is twice as many as the number of customer queues, i.e., $m = 2n$. The compatibility graph is given by

$$E = \{(i, j) \in [2n] \times [n] : j \in \{i+k\} \cup \{(i+k-n)^+\}, k = 0, 1\}.$$

The demand and supply curve are assumed to be

$$F_j(\lambda_j) = 6 - \lambda_j, \forall j \in [m] \qquad \text{and} \qquad G_i(\mu_i) = \mu_i, \forall i \in [n],$$

respectively. The parameters of pricing policy are similar to the previous case when $m = n$. We report the profit loss for $\eta \in \{10, 100, 500, 1000, 2000, 5000, 10000\}$ when $m = 8$ and $n = 4$ and its associated plot in logarithmic scale in Figure 4.8. The result shows that when η is larger, the profit loss when using the two-price policy grows significantly slower, compared to when using the fluid pricing policy. Moreover, we can observe that in this case the benefit of the max-weight matching policy over the randomized matching policy when combined with any pricing policy is negligible. Figure 4.8 (right) shows that the fitted orders of profit loss with respect to η of FP+MW and FP+Rand are 0.49 and 0.48 and those of TP+MW and TP+Rand are 0.33 and 0.32. This observation confirms the results from Theorem 5 and Theorem 8 as well as Theorem 6 and Theorem 9, which state that the orders of profit loss with respect to η of FP+MW and FP+Rand are 1/2 and those of TP+MW and TP+Rand are 1/3 respectively. Figure 4.9 shows the profit loss for $n \in \{4, 6, 8, \ldots, 20\}$ and $m = 2n$ when $\eta = 10000$ and its associated log-log plot. This figure shows the superiority of the max-weight policy over the randomized matching policy discussed in Section 4.6.2.

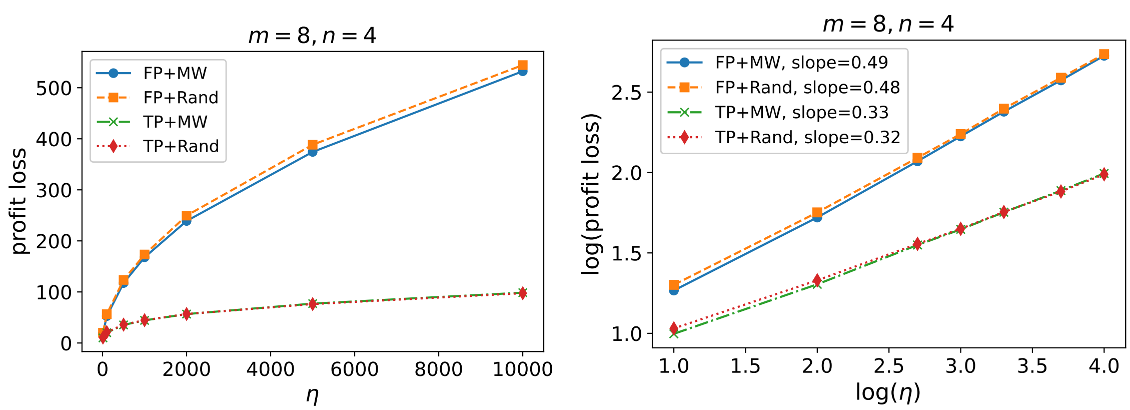


Figure 4.8: Profit losses under FP+MW, FP+Rand, TP+MW and TP+Rand for different η (left) and its associated log-log plot (right) when $m = 8$ and $n = 4$.

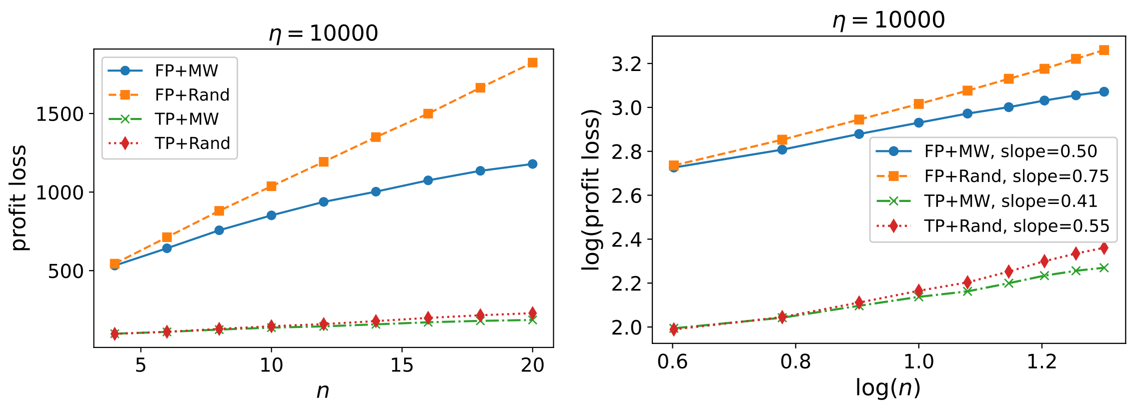


Figure 4.9: Profit losses under FP+MW, FP+Rand, TP+MW and TP+Rand for different n (left) and its associated log-log plot (right) when $\eta = 10000$.

4.8 Conclusion

In this chapter, we present a model of dynamic pricing and matching for two-sided queueing systems. The system is formulated as a Markov decision process, and a fluid approximation model is considered. We presented a fluid pricing and max-weight matching policy and showed that it achieves $O(\sqrt{\eta})$ optimality rate. Furthermore, we proposed a dynamic pricing and max-weight policy, which achieves $O(\eta^{1/3})$ optimality rate. We also show that this scaling of $O(\eta^{1/3})$ matches the lower bound for a broad family of policies. We also demonstrate the advantage of max-weight matching over randomized matching. Under the complete resource pooling condition, we show that max-weight matching achieves $O(\sqrt{n})$ and $O(n^{1/3})$ optimality rates for static and two-price policies, respectively, where n is the number of customer and server types. In comparison, the randomized matching policy may have an $\Omega(n)$ optimality rate.

CHAPTER 5
NEURAL NETWORK CHOICE MODEL

5.1 Introduction

One problem that often arises for business, e.g. e-commerce, hotel, airline, is to decide which subset of products to offer to customers. The subset of products offered to customers is called an assortment. The goal of the firm is to, for example, decide which assortment to offer to maximize the expected revenue.

One important ingredient for the firm to answer this problem is to understand customer preference. Historically, independent demand model, which a customer prefers a particular product regardless of which assortment is offered, is commonly assumed. However, this is an unrealistic assumption, because it ignores product substitution effect. Therefore, choice-based demand model is considered. Discrete choice model is known as one of essential tools to model individual behavior. It describes decision makers' choice when a different subset of products is offered.

One of the most popular choice models belongs to this class is the multinomial logit (MNL) model (McFadden, 1974). The main advantage of MNL is its simplicity in terms of estimation and interpretation. However, the MNL is derived from one restrictive assumption, i.e., independently and identically distributed (IID) error terms. This results in a well-known Independence of Irrelevant Alternatives (IIA) property or Luce's Axiom of Choice (Luce, 1959). This property implies that the relative likelihood of choosing any two alternatives from the choice set is independent of the availability of other alternatives in the choice set. This IIA property seems to be too restrictive and may be violated in reality. To overcome this shortfall, researchers have come up with a variety of other discrete choice models, which relax this restrictive assumption, such as nested logit (NL) model (Domencich and McFadden, 1975) , mixed logit model (McFadden and Train, 2000; Hensher and Greene, 2002) and latent class logit model (Greene and Hensher, 2003).

All choice models mentioned earlier is derived from random utility maximization (RUM) principle, where customers are assumed to choose items which have the highest utility among the offered products. The assumption that choice is based on utility maximization implies regularity property, i.e., adding an option to a choice set cannot increase the choice probability for any of the original choice options. This phenomenon can be observed in wide range of situations. However, in reality customers might not act rationally; therefore, the regularity property might no longer hold; see, e.g., Huber, Payne, and Puto (1982) and Tversky, Slovic, and Kahneman (1990). In this case, any model, no matter how complex it is, in this RUM class cannot explain this behavior. Recently, Jagabathula and Rusmevichientong (2019) are able to identify when RUM model is not suitable for the dataset in predicting the choice behavior through the limit of rationality (LoR), which is the degree of inconsistency between the observed data and RUM model.

Because some choice behaviors cannot be explained by models which derive from RUM principle, there are several efforts in explaining the behavior beyond RUM class. Some examples include the general attraction model (GAM) (Gallego, Ratliff, and Shebalov, 2015), the generalized Luce model (Echenique and Saito, 2018), and the perception adjusted Luce model (PALM) (Echenique, Saito, and Tserenjigmid, 2018). However, these choice models do not subsume the RUM class. Therefore, they potentially perform worse than the RUM class. Berbeglia (2018) propose the generalized stochastic preference (GSP) model, which subsumes the RUM class and can explain some choice phenomena that are outside the RUM class. Nevertheless, there exists choice behaviors which cannot be explained by this model. This motivates the study of more general choice model, which is able to explain any choice behavior.

Due to the success of machine learning (ML) method in extracting complex behaviors, some researchers apply machine learning techniques to model general choice behavior. For instance, Osogami and Otsuka (2014) introduce the restricted Boltzmann machine (RBM) choice model, which extends the MNL model. The RBM choice model can capture irrational behaviors and can be trained easily using existing algorithms for RMB. Moreover, Chen and Mišić (2019) and Chen, Gallego, and Tang (2021) uses a set of decision trees (decision forest) to model customer choice. They show that any type of choice behavior can be modeled a probability distribution over a set of decision trees.

This chapter aims to develop a neural network framework which is able to explain customer choice behavior. We investigate the performance of our proposed model and compare with other well-known discrete choice models through extensive numerical experiments using both synthetic dataset and real world datasets. We show that our proposed model, which builds upon neural network structure, consistently outperforms the other models in all datasets.

5.1.1 Notation

For a positive integer m, let $[m]$ denote the set $\{1,2,\ldots,m\}$. For a set $\mathscr{S}$, let $|\mathscr{S}|$ denote the cardinality of $\mathscr{S}$. Suppose v and w are vectors of dimension m and n respectively. We denote the concatenated vector of dimension $m+n$ by (v,w). Whenever we refer to a vector, we mean a column vector. Let $\mathbb{1}_{\{\cdot\}}$ denote an indicator function.

5.2 Model

Let n be the total number of product, which includes a no-purchase or outside optionand and $\mathscr{J} = [n]$ denote a set of all products. An assortment $\mathscr{A}$ is a subset of products, i.e., $\mathscr{A} \subseteq \mathscr{J}$. When an assortment $\mathscr{A}$ is offered, the customer can choose any item j in that assortment, that is, $j \in \mathscr{A}$. We want to study the discrete choice model, which is a conditional probability $\mathbb{P}(\cdot|\cdot) : \mathscr{J} \times 2^{\mathscr{J}} \to [0,1]$. Specifically, $\mathbb{P}(j|\mathscr{A})$ denotes the probability that an item $j \in \mathscr{A}$ is chosen from an assortment $\mathscr{A} \subseteq \mathscr{J}$. We note that the choice model must satisfy the following properties:

- $\mathbb{P}(j|\mathscr{A}) \geq 0$ for all $j \in \mathscr{J}$ and $\mathscr{A} \subseteq \mathscr{J}$,
- $\mathbb{P}(j|\mathscr{A}) = 0$ for all $j \notin \mathscr{A}$ and $\mathscr{A} \subseteq \mathscr{J}$,
- $\sum_{j \in \mathscr{A}} \mathbb{P}(j|\mathscr{A}) = 1$ for all $\mathscr{A} \subseteq \mathscr{J}$.

We will use a neural network architecture to model choice behavior. A neural network aims to approximate an output as some complex function of an input. It generally consists of an input layer $\mathscr{I}$, optional hidden layer(s) $\mathscr{H}^k$ for $k \in [K]$, where K is the number of hidden layers and k is the kth hidden layer, and an output layer $\mathscr{O}$. Each node in hidden layers and output layers can do non-linear transformation of the input through the activation function. This allows the network to learn and model more complex behaviors.

The input layer represents an offered assortment. It consists of n nodes, i.e., $|\mathscr{I}| = n$, where n is the number of products. Each node $i \in \mathscr{I}$ takes a binary variable which equals to 1 when product i is in an assortment $\mathscr{A}$ and 0 otherwise. Each node in hidden layers possesses an activation function. The number of hidden layers, K, and the number of nodes in each hidden layer, $|\mathscr{H}^k|$ for $k \in [K]$ are hyperparemeters which can arbitrarily set. Our network has one more layer called utility layer $\mathscr{U}$, which is an intermediate layer before producing probability in the output layer. The number of nodes in this utility layer equals to the number of items n, i.e., $|\mathscr{U}| = n$. The utility layer represents the utility or score of each product. The last layer is the output layer $\mathscr{O}$, which consists of n nodes, that is, $|\mathscr{O}| = n$. This layer represents the probability of selecting each product in an assortment. Note that nodes, which correspond to the products which are not offered in assortment, have value of zero in the output layer.

The input layer, the hidden layers and the utility layer are fully connected, e.g., all the nodes in one layer are connected to every node in a subsequent layer. Let $W^{(\mathscr{I})}$ and $b^{(\mathscr{I})}$ be a weight vector and a bias vector of an input layer respectively. Moreover, we let $W^{(\mathscr{H}^k)}$ and $b^{(\mathscr{H}^k)}$ denote a weight vector and a bias vector of a hidden layer k for $k \in [K]$. Let $a^{(\mathscr{H}^k)}$ denote an activation function of a hidden layer k for $k \in [K]$. Suppose x is an input vector of binary variables of dimension n, i.e., $x \in \{0,1\}^n$, and u is a real value vector of dimension n, i.e., $u \in \mathbb{R}^n$, in the utility layer. Note that, for a product $j \in \mathscr{J}$ and an offered assortment $\mathscr{A} \subseteq \mathscr{J}$, we have

$$x_j = \begin{cases} 1, & \text{if } j \in \mathscr{A} \\ 0, & \text{otherwise} \end{cases}.$$

Formally, we can write a function mapping the input x to the intermediate output u as

$$\begin{aligned} q^{(\mathscr{H}^1)} &= a^{(\mathscr{H}^1)}\left(W^{(\mathscr{I})\top}x + b^{(\mathscr{I})}\right) \\ q^{(\mathscr{H}^2)} &= a^{(\mathscr{H}^2)}\left(W^{(\mathscr{H}^1)\top}q^{(\mathscr{H}^1)} + b^{(\mathscr{H}^1)}\right) \\ &\;\;\vdots \\ u &= a^{(\mathscr{H}^K)}\left(W^{(\mathscr{H}^K)\top}q^{(\mathscr{H}^{K-1})} + b^{(\mathscr{H}^K)}\right), \end{aligned}$$

where $q^{(\mathscr{H}^k)}$ is the output of the kth hidden layer for $k \in [K]$.

Unlike a normal multi-class neural network, we do not apply softmax function directly to the utility layer to get choice probability of each product. We instead apply softmax function only to items which appear in the assortment. Suppose p is a probability output vector of dimension n, i.e., $p \in [0,1]^n$. We can write, for $j \in [n]$,

$$p_j = \frac{e^{u_j}\,\mathbb{1}_{\{x_j=1\}}}{\sum_{\ell=1}^{n} e^{u_\ell}\,\mathbb{1}_{\{x_\ell=1\}}}.$$

This is to ensure that the probability of choosing product outside an offered assortment is zero.

Figure 5.1 shows an example of neural network architecture for modeling choice behavior when the total number of products is five, the number of hidden layers is two, and the number of nodes in each hidden layer is ten.

Suppose $\mathscr{T}^{\text{train}}$ is a collection of customers in training set. For each customer $t \in \mathscr{T}^{\text{train}}$,

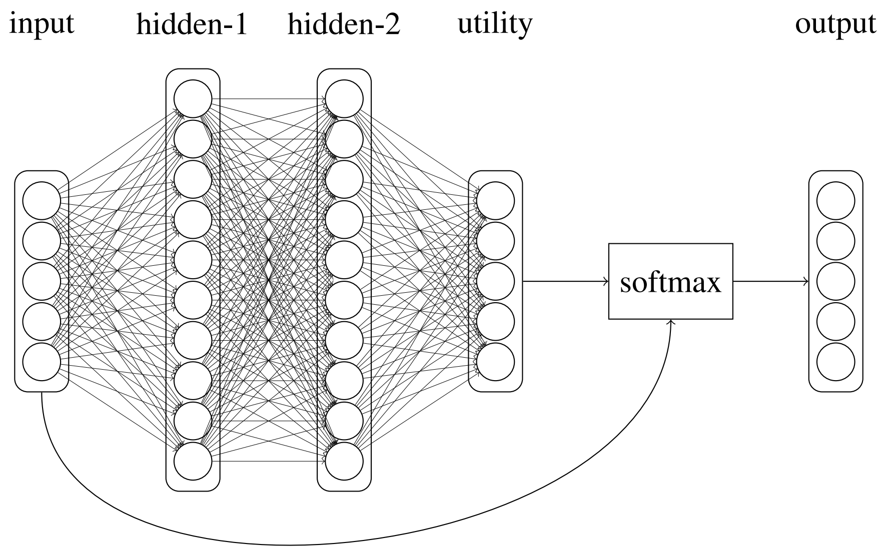


Figure 5.1: An example of neural network architecture for choice modeling when there are five products and two hidden layers whose number of nodes is ten.

we let x_t be an input vector of binary variables of dimension n, $x_t \in \{0,1\}^n$, associated with an offered assortment $\mathscr{A}_t \subseteq \mathscr{J}$. Specifically, $x_{t,j} = 1$ if $j \in \mathscr{A}_t$ and $x_{t,j} = 0$ otherwise. Let p_t denote a probability output vector from the model of dimension n of customer t. To train the model, we consider cross-entropy loss function. The cross-entropy measures the performance of classification model, whose output is a probability distribution. Specifically, it measures the difference between two probability distributions, the true and estimated distributions. For customer t, we let y_t be a unit vector of dimension n represents an actual choice, i.e., $y_{t,j} = 1$ if item j is chosen by customer t and $y_{t,j} = 0$ otherwise. Let $\mathbb{P}$ and $\hat{\mathbb{P}}$ be the true distribution and the estimated distribution respectively. Suppose Y is a random variable of item selected. The cross-entropy loss function is defined as

$$\begin{aligned} H(\mathbb{P},\hat{\mathbb{P}}) &= -\frac{1}{|\mathscr{T}^{\text{train}}|}\sum_{t\in\mathscr{T}^{\text{train}}}\sum_{j\in\mathscr{A}_t}\mathbb{P}(Y=j|\mathscr{A}_t)\log\hat{\mathbb{P}}(Y=j|\mathscr{A}_t) \\ &= -\frac{1}{|\mathscr{T}^{\text{train}}|}\sum_{t\in\mathscr{T}^{\text{train}}}\sum_{j\in\mathscr{A}_t} y_{t,j}\log p_{t,j}. \end{aligned}$$

It can be seen that, for each $t \in \mathscr{T}^{\text{train}}$, when the estimated model produces the probability of the right class as 1, the cross-entropy term of customer t equals to 0. Moreover, this term grows, as the produced probability of the right class decreases. That is, when the cross-entropy loss is small, the estimated model is close to the true model. The training procedure aims to minimize this cross-entropy loss. In other words, it tries to get the estimated model which resembles the true model. In fact, it is easy to verify that this problem is equivalent to maximum log-likelihood estimation.

5.2.1 Model with Features

We can extend the neural network choice model we discussed earlier to incorporate observed choice characteristics. In this case, the input vector x is no longer a binary vector and its dimension of input vector is no longer n. Suppose R is the number of observed features after feature engineering. The input vector is split into $R+1$ parts of equal length, which equals to the number of products n. The vector $x^{(0)}$ represents an offered assortment, similar to the input vector in no feature case. Particularly, $x^{(0)}$ is a binary vector of dimension n, and $x_j^{(0)}$ equals to 1 when product j is in an assortment $\mathscr{A}$ and 0 otherwise. For $r \in R$, the vector $x^{(r)}$ is a vector corresponds to feature r, where each element $x_j^{(r)}$ describes feature r of product j. For $j \notin \mathscr{A}$, we set $x_j^{(r)} = 0$. Let $x_{scaled}^{(r)}$ denote the min-max normalized vector of $x^{(r)}$ for $r \in [R]$. Note that we have $x_{scaled}^{(r)} \in [0,1]^n$ for all $r \in [R]$ in training set. The input vector x is then constructed as $x = (x^{(0)}, x_{scaled}^{(1)}, \ldots, x_{scaled}^{(R)})$, where $x^{(0)}, x_{scaled}^{(1)}, \ldots, x_{scaled}^{(R)}$ are vectors of dimension n. Therefore, the input vector x has dimension of $(R+1)n$, whose value is between 0 and 1 inclusive, i.e., $x \in [0,1]^{(R+1)n}$.

5.3 Numerical experiment

In this section, we will compare the effectiveness of different discrete choice models through numerical experiments. To quantify the effectiveness of the models, we test the out-of-sample predictive ability. Specifically, we randomly divide the data into training data and testing data of equal size. The training data is used to estimate the model, and the testing data is then used to measure the predictive power of the models. Suppose $\mathscr{T}^{\text{test}}$ is a collection of customers in testing set. The metric we use is cross-entropy which is defined as

$$H(\mathbb{P},\hat{\mathbb{P}}) = -\frac{1}{|\mathscr{T}^{\text{test}}|}\sum_{t\in\mathscr{T}^{\text{test}}}\sum_{j\in\mathscr{A}_t} y_{t,j}\log p_{t,j}.$$

Recall that the model which has smaller cross-entropy means it has better predicting power. The empirical cross-entropy is given by

$$H(\hat{\mathbb{P}},\hat{\mathbb{P}}) = -\frac{1}{|\mathscr{T}^{\text{test}}|}\sum_{t\in\mathscr{T}^{\text{test}}}\sum_{j\in\mathscr{A}_t} p_{t,j}\log p_{t,j}.$$

It can be shown by Jensen's inequality that the empirical cross-entropy is the smallest achievable cross-entropy any model can be obtained.

The models we consider include

1. The neural network choice model discussed in Section 5.2. We note that the number of hidden layers, the number of nodes in each hidden layer and an activation function of each node in each hidden layer can be arbitrarily set. For simplicity, we let each hidden layer have the same number of nodes, which equal to a multiple of the number of nodes in the input layer. Specifically, the number of nodes in each hidden layer $|\mathscr{H}^k| = cn$ for $k \in [K]$, where $c \in \mathbb{N}$ is a multiplier of the number of nodes in the input layer. We assume that each node has the same activation function, which is a leaky rectified linear unit (leaky ReLU) (Maas, Hannun, and Ng, 2013). A leaky ReLU is defined as

$$a(x) = \begin{cases} \alpha x, & \text{if } x < 0 \\ x, & \text{otherwise} \end{cases},$$

where α is a hyperparameter. We consider when the number of hidden layers K is 0,1 and 2, namely

(a) the neural netwok choice model with zero hidden layer (NN-0),

(b) the neural netwok choice model with one hidden layer (NN-1),

(c) the neural netwok choice model with two hidden layers (NN-2).

We test the model when the number of nodes in each hidden layer is $1, 2, \ldots, 10$ times of the number of products. For the activation function of each node in the hidden layers, we use the Leaky ReLU activation function with parameter $\alpha = 0.1$.

We use TensorFlow Keras to train the neural network choice model. Adam optimization with default parameters is used. We use early stopping to avoid overfitting. We let the procedure terminate if the cross-entropy of validation set, which accounts for 20% of training set, does not decrease by 0.001 for 8 consecutive epochs.

2. The multinomial logit (MNL) model. In this model, a customer chooses product $j \in \mathscr{A}$ when an assortment $\mathscr{A}$ is offered with probability

$$\mathbb{P}(Y = j | \mathscr{A}) = \frac{e^{u_j}}{\sum_{\ell \in \mathscr{A}} e^{u_\ell}},$$

where parameter u_j is utility of product j for $j \in \mathscr{J}$. The parameters are obtained from maximum likelihood estimation.

3. The restricted Boltzmann Machine (RBM) choice model proposed by Osogami and Otsuka (2014). The model extends the MNL model. Specifically, a customer chooses product $j \in \mathscr{A}$ when an assortment $\mathscr{A}$ is offered with probability

$$\mathbb{P}(Y = j | \mathscr{A}) = \frac{e^{u_j} \prod_{m \in \mathscr{M}} (T_{\mathscr{A}}^m + V_j^m)}{\sum_{\ell \in \mathscr{A}} e^{u_\ell} \prod_{m \in \mathscr{M}} (T_{\mathscr{A}}^m + V_\ell^m)},$$

where $T_{\mathscr{A}}^m := \sum_{\ell \in \mathscr{A}} T_\ell^m$. Note that the model has parameters $u_j, T_{\mathscr{A}}^m$ and V_j^m for $j \in \mathscr{A}$

and $m \in \mathscr{M}$, where $\mathscr{M}$ is a set of indices. They show that this choice model can be represented by RBM. So, we use existing algorithm for RBMs to estimate parameters. Specifically, we use sklearn.neural_network.BernoulliRBM with learning rate 0.01 for 100 epochs to learn parameters.

4. The spiked multinomial logit (SMNL) model proposed by Cao, Kleywegt, and Wang (2019). Let c denote the cheapest product in an assortment $\mathscr{A}$. In this model, a customer given an assortment $\mathscr{A}$ chooses product $j \in \mathscr{A}$ with probability

$$\mathbb{P}(Y = j|\mathscr{A}) = \frac{e^{u_j}(1 - \mathbb{1}_{\{c=j\}}) + e^{v_j}\,\mathbb{1}_{\{c=j\}}}{\sum_{\ell \in \mathscr{A}}\left[e^{u_\ell}(1 - \mathbb{1}_{\{c=\ell\}}) + e^{v_\ell}\,\mathbb{1}_{\{c_t=\ell\}}\right]},$$

where u_j is a regular utility and v_j is a special utility when product j is the cheapest product in an assortment $\mathscr{A}$ for all $j \in \mathscr{J}$. The parameters are obtained from maximum likelihood estimation.

5. The general attraction model (GAM) proposed by Gallego, Ratliff, and Shebalov (2015). Let $j_0 \in \mathscr{J}$ denote a no-purchase option. In this model, a customer given an assortment $\mathscr{A}$ chooses product $j \in \mathscr{A}$ with probability

$$\mathbb{P}(Y = j|\mathscr{A}) = \begin{cases} \frac{e^{u_j}}{\sum_{\ell \in \mathscr{A}} e^{u_\ell} + \sum_{\ell \in \mathscr{J} \setminus \mathscr{A}} e^{v_\ell}}, & \text{if } j \neq j_0 \\ \frac{e^{u_{j_0}} + \sum_{\ell \in \mathscr{J} \setminus \mathscr{A}} e^{v_\ell}}{\sum_{\ell \in \mathscr{A}} e^{u_\ell} + \sum_{\ell \in \mathscr{J} \setminus \mathscr{A}} e^{v_\ell}}, & \text{if } j = j_0 \end{cases},$$

where u_j is a regular utility and v_j is a shadow utility when product j is not offered in an assortment. The model is fit by maximum likelihood estimation.

5.3.1 Synthetic Data

Dataset

We follow Cao, Kleywegt, and Wang (2019) for synthetic data generation process to obtain data which has a spike phenomenon. Specifically, there are p types of customer whose arrival rates are all equal. Each type of customer considers different subset of products and will choose the cheapest product in her consideration set that is available in an offered assortment. If no product in her consideration set available, the customer will choose no-purchase option.

Suppose product 1 refers to a no-purchase option. The remaining $n+1$ types of products are ranked from high to low price. Specifically, product 2 has the highest price, and product n has the lowest price. Each type of customer $\ell \in [m]$ considers product only in specific range $[\underline{i}_\ell, \bar{i}_\ell]$. Let $U\{\underline{u}, \bar{u}\}$ denote the discrete uniform distribution which can have integer values from $\underline{u}$ and $\bar{u}$. For each type of customer $\ell \in [m]$, $\underline{i}_\ell$ is drawn from $U\{2, n\}$, and then $\bar{i}_\ell$ is drawn from $U\{\underline{i}_\ell, n\}$. If the range $[\underline{i}_\ell, \bar{i}_\ell]$ coincides with the range of any preceding type of customer, we keep regenerating until the range is unique. Let τ be the number of customers. For each $t \in \tau$, assortment $\mathscr{A}_t$ is $[n_t]$ where n_t is drawn from $U\{2, n\}$. The type of customer arrives at t is drawn from $U\{1, p\}$.

We consider $n = 15$ and $p = 30$. We generate 3000 customers for training set, and another 3000 customers for testing set.

Result

Figure 5.2 illustrates cross-entropy of the MNL, SMNL, GAM, NN-0, NN-1, NN-2 and RBM models for training set (left) and testing set (right) when the number of nodes in each hidden layer is 1,2,...,10 times the number of products. It can be seen that the performances of each model are consistent in both training and testing set. That is, the RMB choice model is the worst performer. the GAM and MNL models have roughly the same level of cross-entropy, but the GAM performs slightly better than the MNL models. The SMNL model does significantly better, because the data we consider has a spike phenomenon which the SMNL model can capture. However, the neural network choice models, NN-0, NN-1 and NN-2, can do better than the SMNL model. This is an interesting result, as the SMNL model requires the additional knowledge of price ranking of products,

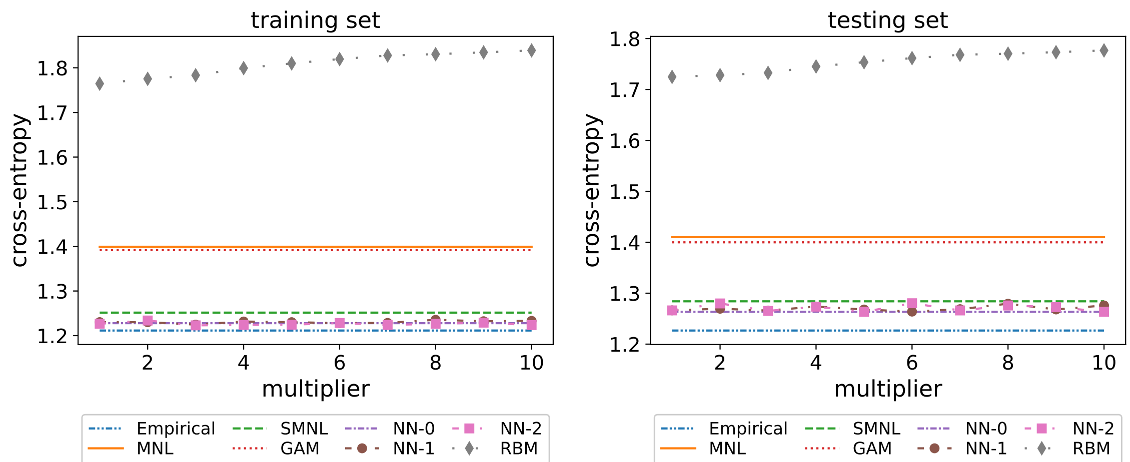


Figure 5.2: Cross-entropy of the MNL, SMNL, GAM, NN-0, NN-1, NN-2 and RBM models for training set (left) and testing set (right) of synthetic data.

while the neural network models (and others) do not. This result shows how powerful the neural network choice models are in predicting underlying choice behavior. Moreover, we can observe that the RMB choice model does not perform well as the number of nodes in the hidden layer increases in both training and testing sets. As we fix the number of epochs when we train the RMB choice model, this might happen because of under-training in more complex model. However, we cannot observe such trend in the NN-1 and NN-2 models. Moreover, for this synthetic dataset, it is inconclusive whether the number of hidden layers helps improve the performance of the neural network models.

Table 5.1 summarizes the improvement of cross-entropy of the neural network models, NN-0, NN-1 and NN-2, over the MNL and SML models for synthetic testing data. It can be seen that the neural network models outperform the MNL and SMNL models by approximately 10.07% and 1.25% respectively.

5.3.2 Hotel Data

Dataset

We apply each model to "Hotel 1" in publicly available hotel data found in Bodea, Ferguson, and Garrow (2009). The data includes transient customers whose check-in dates be-

Table 5.1: Improvement of cross-entropy of the NN-0, NN-1 and NN-2 models over the MNL and SMNL models for synthetic testing data.

Multiplier	Improvement over MNL			Improvement over SMNL		
	NN-0	NN-1	NN-2	NN-0	NN-1	NN-2
1	10.39%	10.20%	10.20%	1.60%	1.38%	1.39%
2	10.39%	9.97%	9.23%	1.60%	1.13%	0.32%
3	10.39%	10.23%	10.28%	1.60%	1.41%	1.47%
4	10.39%	9.67%	9.68%	1.60%	0.80%	0.82%
5	10.39%	10.08%	10.40%	1.60%	1.26%	1.61%
6	10.39%	10.37%	9.21%	1.60%	1.57%	0.29%
7	10.39%	10.03%	10.21%	1.60%	1.20%	1.40%
8	10.39%	9.27%	9.52%	1.60%	0.36%	0.64%
9	10.39%	10.07%	9.75%	1.60%	1.24%	0.89%
10	10.39%	9.52%	10.40%	1.60%	0.64%	1.60%

tween March 12, 2007, and April 15, 2007, in one of the continental U.S. hotels. We only consider bookings which associate with product availability information, i.e., the merge indicator equals to 1. We define product as a combination of room type and rate code. Each booking record corresponds to one customer. An assortment offered to each customer is defined as all available products associated with such booking record plus a no-purchase option. For each booking record, if the purchased product indicator equals to one, it means customer chooses such product in an assortment. However, if all purchased product indicators equal to zero, it means customer chooses no-purchase option. As we need price ranking of products to estimate the SMNL model, we use an average arrival rate of each product as a proxy of its price.

After data preprocessing, "Hotel 1" data has n=81, 191 different assortments and 1595 booking records (customers). We randomly select 50% of booking records, which is equivalent to 797 entries, to construct training set and the remaining 798 entries go to testing set.

Result

Figure 5.3 shows cross-entropy of the MNL, SMNL, GAM, NN-0, NN-1, NN-2 and RBM models for training set (left) and testing set (right) when the number of nodes in each hidden layer is 1,2,...,10 times the number of products. We observe that the neural network models generally outperforms other models in both training and testing sets, with the exception of the NN-2 model with multiplicative factors of 5, 8, 9 and 10. For the neural network models, we can observe that more complex models fit data better, in particular, the NN-2 model fits data better than the NN-1 model, and the NN-1 model fits data better than the NN-0 model. In training set, we also observe that when the number of nodes in hidden layer increases, both the NN-1 and NN-2 models tend to have smaller cross-entropy. However, when it comes to predicting power, more complex models performs worse. Moreover, we can see large discrepancies between cross-entropy of training and test sets. These observations might be caused by model over-fitting, since the number of training data is small comparing to the number of products and the number of assortments. Moreover, we can see that the predicting power of the MNL model, which belongs to the RUM class, is better than the GAM and SML models, which can model behaviors beyond the RUM class. This result is consistent with what Jagabathula and Rusmevichientong (2019) mention. Table

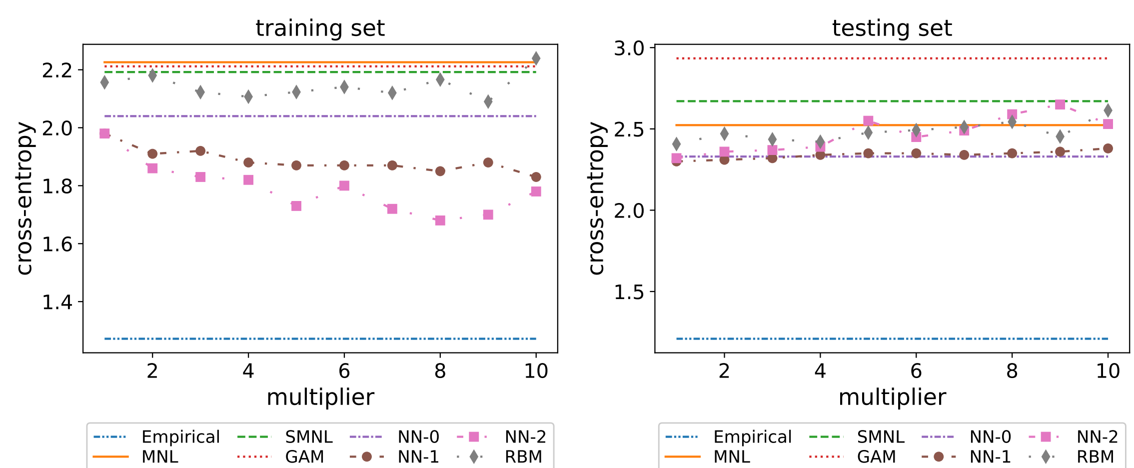


Figure 5.3: Cross-entropy of the MNL, SMNL, GAM, NN-0, NN-1, NN-2 and RBM models for training set (left) and testing set (right) of hotel data.

5.2 summarizes the improvement of cross-entropy of the neural network models, NN-0, NN-1 and NN-2, over the MNL model for hotel testing data. It can be seen that majority of the neural network configurations outperform the MNL model.

Table 5.2: Improvement of cross-entropy of the NN-0, NN-1 and NN-2 models over the MNL model for hotel testing data.

Multiplier	Improvement over MNL		
	NN-0	NN-1	NN-2
1	7.61%	8.76%	8.06%
2	7.61%	8.50%	6.63%
3	7.61%	8.15%	6.08%
4	7.61%	7.22%	5.35%
5	7.61%	6.84%	-1.11%
6	7.61%	6.88%	2.94%
7	7.61%	7.11%	1.35%
8	7.61%	6.69%	-2.82%
9	7.61%	6.42%	-5.01%
10	7.61%	5.65%	-0.21%

5.3.3 IRI Data

Dataset

We focus on IRI Academic Dataset (Bronnenberg, Kruger, and Mela, 2008). This dataset includes purchase transactions of consumer packaged goods for both grocery stores and drug stores over 47 markets across the US. We note that the same dataset is used by Jagabathula and Rusmevichientong (2019) and Chen and Mišić (2019).

Due to the large volume of dataset, we consider the data from the first four weeks of 2007 from drug stores. Specifically, we focus on "milk" product category. We define product as the items with the same vendor code, which can be found from digits 4 to 8 in the universal product code (UPC). We focus on the top 30 purchased vendor codes and we aggregate the remaining purchased transaction as the 31^{th} product. Therefore, we have $n = 31$. The aggregation by vendor code is a commonly-used technique for data preprocessing; see Bronnenberg and Mela (2004) and Nijs, Srinivasan, and Pauwels (2007).

An assortment offered to each customer is defined as all products, which is purchased at least once, during a week and at a particular store. As we need price ranking of products to estimate the SMNL model, we use an average price per unit of all items with the same vendor code as a proxy of the price.

5.3.4 Result

Figure 5.4 shows cross-entropy of the MNL, SMNL, GAM, NN-0, NN-1, NN-2 and RBM modles for training set (left) and testing set (right) when the number of nodes in each hidden layer is 1,2,...,10 times the number of products. We observe that the results from training set and testing set are consistent. Specifically, the RBM choice model is the worst performer. The MNL and SMNL models have roughly similar performance, while the GAM does slightly better. The neural network, NN-0, NN-1 and NN-2, models perform well, as we observe some improvement gaps from the other choice models.

After data preprocessing, "milk" data has n=31, 108 different assortments and 21864 transactions. We randomly select 50% of booking records, which is equivalent to 10932 entries, to construct training set and the remaining 10932 entries go to testing set. Table

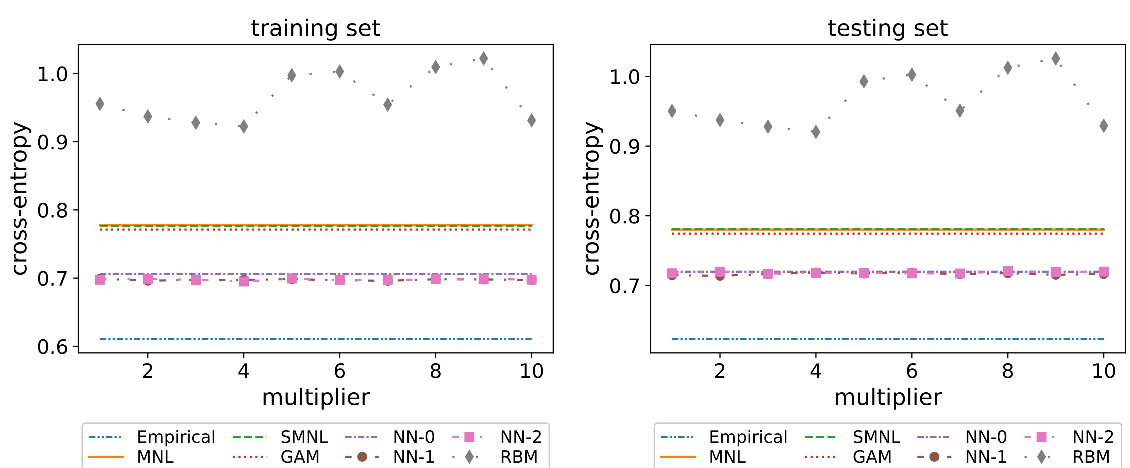


Figure 5.4: Cross-entropy of the MNL, SMNL, GAM, NN-0, NN-1, NN-2 and RBM models for training set (left) and testing set (right) of milk category from IRI Academic Dataset.

5.3 summarizes the improvement of cross-entropy of the neural network models, NN-0, NN-1 and NN-2, over the MNL model for milk testing data. It can be seen that the neural network configurations outperform the MNL model by 7.94% on average. It also shows that the model with no hidden layer has worse performance than one with hidden layer(s). This result suggests that when we have sufficiently large dataset for training, we can avoid over-fitting in more complex models.

Table 5.3: Improvement of cross-entropy of the NN-0, NN-1 and NN-2 models over the MNL model for milk testing data.

Multiplier	Improvement over MNL		
	NN-0	NN-1	NN-2
1	7.74%	8.39%	8.04%
2	7.74%	8.50%	7.74%
3	7.74%	8.18%	8.12%
4	7.74%	7.91%	7.96%
5	7.74%	8.08%	7.97%
6	7.74%	7.92%	8.06%
7	7.74%	8.19%	8.08%
8	7.74%	8.01%	7.65%
9	7.74%	8.31%	7.81%
10	7.74%	8.17%	7.74%

5.4 Conclusion

In this chapter, we propose the neural network choice model, which build upon neural network framework, to model general choice behaviors, either within RUM class or beyond RUM class. Numerical experiments show that our proposed models consistently outperform other well-known discrete choice models, which belong to RUM class or beyond RUM class, in both synthetic and real world datasets.

Appendices

APPENDIX A
RE-SOLVING HEURISTIC WITH UNIFORMLY BOUNDED LOSS FOR NETWORK REVENUE MANAGEMENT

A.1 Proofs for Section 2.4

In this section, we provide complete proofs for the results on the IRT policy in Section 2.4.

A.1.1 Proof of Theorem 1

Proof of Theorem 1. Given remaining capacity $C(t_1)$ at time $t_1 \in [0,T]$, let $x(t_1)$ be an optimal solution to the following LP

$$\max_x \Big\{ \sum_{j=1}^n r_j x_j \;\Big|\; \sum_{j=1}^n A_j x_j \le C(t_1)/(T-t_1), \text{ and } 0 \le x_j \le \lambda_j, \forall j \in [n] \Big\}.$$

For any $t_2 \in (t_1, T]$, let $V^{\mathsf{SPA}}(t_1,t_2)$ denote the revenue earned in $[t_1,t_2)$ under a static probabilistic allocation policy, where class j customers are accepted with probability $x_j(t_1)/\lambda_j$. Let $V^{\mathsf{SPA}'}(t_1,t_2)$ be the revenue earned in $[t_1,t_2)$, where a class j customer is accepted with the following probability:

- 0, if $x_j(t_1) < \lambda_j (T-t_1)^{-1/4}$
- 1, if $x_j(t_1) > \lambda_j (1-(T-t_1)^{-1/4})$
- $x_j(t_1)/\lambda_j$, otherwise.

Let $V^{\mathrm{HO}}(t_1,T)$ denote the revenue earned from solving the hindsight optimum in $[t_1,T]$. That is, $V^{\mathrm{HO}}(t_1,T)$ is the optimal revenue given the remaining capacity at t_1 and a sample path of demand in $(t_1, T]$, given by

$$V^{\text{HO}}(t_1,T) = \max_y \Big\{ \sum_{j=1}^n r_j y_j \;\Big|\; \sum_{j=1}^n A_j y_j \le C(t_1), \text{ and } 0 \le y_j \le \Lambda_j(T) - \Lambda_j(t_1), \forall j \in [n] \Big\}.$$

Consider policy IRT^2, which re-solves at $t_1^* = T - T^{5/6}$, and re-solves again at $t_2^* = T - T^{(5/6)^2} = T - T^{25/36}$. Let v^{IRT^2} be the expected revenue of IRT^2. The regret of this policy can be decomposed as

$$\begin{aligned} v^{\text{HO}} - v^{\mathsf{IRT}^2} &= \mathbb{E}[V^{\text{HO}} - V^{\mathsf{IRT}^2}] \\ &= \mathbb{E}[V^{\text{HO}} - V^{\mathsf{HO}^1}] + \mathbb{E}[V^{\mathsf{HO}^1} - V^{\mathsf{HO}^2}] + \mathbb{E}[V^{\mathsf{HO}^2} - V^{\mathsf{IRT}^2}]. \end{aligned} \tag{A.1}$$

The first term of Eq (A.1) is bounded by $O(Te^{-\kappa T^{1/6}})$ as stated in Proposition 1. For the second term of Eq (A.1), the revenue of policy HO^1 and HO^2 are

$$V^{\mathsf{HO}^1} = V^{\mathsf{SPA}'}(0,t_1^*) + V^{\text{HO}}(t_1^*,T), \quad V^{\mathsf{HO}^2} = V^{\mathsf{SPA}'}(0,t_1^*) + V^{\mathsf{SPA}'}(t_1^*,t_2^*) + V^{\text{HO}}(t_2^*,T).$$

Note that policy HO^1 and HO^2 are exactly the same during time $t \in [0, t_1^*)$, so we have

$$V^{\mathsf{HO}^1} - V^{\mathsf{HO}^2} = V^{\text{HO}}(t_1^*,T) - (V^{\mathsf{SPA}'}(t_1^*,t_2^*) + V^{\text{HO}}(t_2^*,T)).$$

Applying part (1) of Proposition 1 to the remaining problem in $(t_1^*, T]$, which has a horizon length $T - t_1^* = \tau_1 = T^{5/6}$, we get

$$\mathbb{E}[V^{\text{HO}}(t_1^*,T) - (V^{\mathsf{SPA}'}(t_1^*,t_2^*) + V^{\text{HO}}(t_2^*,T))] = O(T^{5/6}e^{-\kappa (T^{5/6})^{1/6}}) = O(T^{5/6}e^{-\kappa T^{5/36}}). \tag{A.2}$$

Because the revenue of IRT^2 can be decomposed as

$$V^{\mathsf{IRT}^2} = V^{\mathsf{SPA}'}(0,t_1^*) + V^{\mathsf{SPA}'}(t_1^*,t_2^*) + V^{\mathsf{SPA}}(t_2^*,T),$$

the last term of (A.1) can be bounded by

$$\begin{aligned}\mathbb{E}[V^{\mathsf{HO}^2} - V^{\mathsf{IRT}^2}] &= \mathbb{E}[V^{\mathsf{HO}}(t_2^*, T) - V^{\mathsf{SPA}}(t_2^*, T)] \\ &= O(\sqrt{T - t_2^*}) \\ &= O(T^{(5/6)^2/2}). \end{aligned} \tag{A.3}$$

Eq (A.3) follows the well-known result that static probabilistic allocation has a regret of $O(\sqrt{k})$ for a problem with horizon length k (see Appendix A.2.2).

Combining (A.2) and (A.3), Eq (A.1) is bounded by

$$v^{\mathsf{HO}} - v^{\mathsf{IRT}^2} = O(Te^{-\kappa T^{1/6}}) + O(T^{5/6}e^{-\kappa T^{5/36}}) + O(T^{25/72}). \tag{A.4}$$

Now, consider policy IRT^3, which follows IRT^2 during $t \in [0, t_3^*)$, but re-solves again at time $t_3^* = T - T^{(5/6)^3}$. By the same decomposition argument, the expected regret is given by

$$\begin{aligned} v^{\mathsf{HO}} - v^{\mathsf{IRT}^3} &= \mathbb{E}[V^{\mathsf{HO}} - V^{\mathsf{HO}^1}] + \mathbb{E}[V^{\mathsf{HO}^1} - V^{\mathsf{HO}^2}] + \mathbb{E}[V^{\mathsf{HO}^2} - V^{\mathsf{HO}^3}] + \mathbb{E}[V^{\mathsf{HO}^3} - V^{\mathsf{IRT}^3}] \\ &= O(Te^{-\kappa T^{1/6}}) + O(T^{5/6}e^{-\kappa T^{5/36}}) + O(T^{25/36}e^{-\kappa T^{25/216}}) + O(T^{(5/6)^3/2}). \end{aligned}$$

Let $K = \left\lceil \frac{\log\log T}{\log(6/5)} \right\rceil$. Note that the policy IRT^K coincides with IRT. By induction, if the decision maker re-solves K times, where the uth re-solving time is $t_u^* = T - T^{(5/6)^u}$, the regret is given by

$$v^{\mathsf{HO}} - v^{\mathsf{IRT}} = v^{\mathsf{HO}} - v^{\mathsf{IRT}^K} = \sum_{u=0}^{K-1} O\left((T^{(5/6)^u} \exp\left(-\kappa T^{(5/6)^u/6}\right)\right) + O(T^{(5/6)^K/2}). \tag{A.5}$$

For the first term of the right hand side of Equation (A.5), using the fact that $T^{(5/6)^K} \leq e$,

we have

$$\begin{aligned}\sum_{u=0}^{K-1} T^{(5/6)^u} \exp\left(-\kappa T^{(5/6)^u/6}\right) &= \sum_{\ell=1}^{K} T^{(5/6)^{K-\ell}} \exp\left(-\kappa T^{(5/6)^{K-\ell}/6}\right) \\ &\leq \sum_{\ell=1}^{K} e^{(6/5)^\ell} \exp\left(-\kappa e^{(6/5)^\ell/6}\right) \\ &\leq \sum_{\ell=1}^{\infty} e^{(6/5)^\ell} \exp\left(-\kappa e^{(6/5)^\ell/6}\right) \\ &\leq \int_0^\infty x \exp\left(-\kappa x^{1/6}\right) = O(1).\end{aligned}$$

Thus, the first term of the right hand side of Equation (A.5) is $O(1)$. By the definition of constant K, we have $T^{(5/6)^K/2} \leq e^{1/2}$. Thus, the second term in (A.5) is also $O(1)$. Therefore, we have $v^{\mathsf{HO}} - v^{\mathsf{IRT}} = O(1)$. In addition, this constant factor is independent of the time horizon T and the capacity vector C. □

A.1.2 Proof of Proposition 1

Throughout this subsection, we focus on policies IRT^1 and HO^1 with only one re-solving at $t_1^* = T - T^{5/6}$. We write $t^* := t_1^*$ for simplicity. Let $\Gamma(T) = \alpha \sum_{j:x_j^*=\lambda_j} |\Lambda_j(T) - \lambda_j T|$, where α is a constant whose value is determined by the BOM matrix $A = (a_{lj})_{l\in[m],j\in[n]}$. More specifically, α is the maximum absolute value of the elements in the inverses of all invertible submatrices of the BOM matrix A. In a special case when all entries of A are either 0 or 1, we have $\alpha \leq \max\{1, m \wedge n - 1\}$. We let $\Delta_j(t)$ be the deviation of the number of arrivals of class j customer from its mean in $(t, T]$, i.e., $\Delta_j(t) = \Lambda_j(T) - \Lambda_j(t) - \lambda_j(T - t)$. Define $\tilde{z}_j(t)$ as the number of class j customers accepted up to time t if the algorithm were allowed to go over the capacity limits. For all $j \in [n]$, we define the following events:

$$E_{1,j} = \left\{(T - t^*)x_j^* - \tilde{z}_j(t^*) + t^* x_j^* \geq \Gamma(T)\right\}, \tag{A.6}$$

$$E_{2,j} = \left\{(T - t^*)(\lambda_j - x_j^*) + \tilde{z}_j(t^*) - t^* x_j^* \geq \Gamma(T) + |\Delta_j(t^*)|\right\}. \tag{A.7}$$

The event E is defined as

$$E = \left(\bigcap_{j: x_j^* \geq \lambda_j T^{-1/4}} E_{1,j} \right) \cap \left(\bigcap_{j: x_j^* \leq \lambda_j (1 - T^{-1/4})} E_{2,j} \right). \tag{A.8}$$

Now we will prove Proposition 1.

Proof of Proposition 1. For all $j : x_j^* < \lambda_j T^{-1/4}$, $\tilde{z}_j(t^*) = 0 \leq T x_j^*$. For all $j : x_j^* \geq \lambda_j T^{-1/4}$, event E in (A.8) implies $\tilde{z}_j(t^*) \leq (T - t^*) x_j^* + t^* x_j^* = T x_j^*$. So, suppose event E holds, the capacity constraints for all resources are satisfied up to period t^*, and we have $z_j(t^*) = \tilde{z}_j(t^*)$.

If $x_j^* < \lambda_j T^{-1/4}$, we have $\bar{z}_j - z_j(t^*) = \bar{z}_j - 0 \geq 0$. (Recall that $\bar{z}_j$ is the solution to the hindsight optimum; see Section 2.2.2.) Otherwise, suppose event E holds, by Lemma 9 in Appendix A.4, we have

$$\begin{aligned} \bar{z}_j - z_j(t^*) &\geq T x_j^* - \Gamma(T) - z_j(t^*) + t^* x_j^* - t^* x_j^* \\ &= (T - t^*) x_j^* - \Gamma(T) - (z_j(t^*) - t^* x_j^*) \geq 0, \end{aligned} \tag{A.9}$$

where (A.9) follows from the condition (A.6) and the fact that $z_j(t^*) = \tilde{z}_j(t^*)$.

Similarly, if $x_j^* > \lambda_j(1 - T^{-1/4})$, we have $z_j(t^*) + \Lambda_j(T) - \Lambda_j(t^*) - \bar{z}_j \geq \Lambda_j(t^*) + \Lambda_j(T) - \Lambda_j(t^*) - \Lambda_j(T) = 0$. Otherwise, suppose event E holds, by Lemma 9, we have

$$\begin{aligned} z_j(t^*) + \Lambda_j(T) - \Lambda_j(t^*) - \bar{z}_j &= z_j(t^*) + (T - t^*)\lambda_j + \Delta_j(t^*) - \bar{z}_j \\ &\geq z_j(t^*) + (T - t^*)\lambda_j + \Delta_j(t^*) - T x_j^* - \Gamma(T) \\ &= (z_j(t^*) - t^* x_j^*) + (T - t^*)(\lambda_j - x_j^*) + \Delta_j(t^*) - \Gamma(T) \\ &\geq (T - t^*)(\lambda_j - x_j^*) + (z_j(t^*) - t^* x_j^*) - |\Delta_j(t^*)| - \Gamma(T) \\ &\geq 0, \end{aligned} \tag{A.10}$$

where (A.10) follows from the condition (A.7) and the fact that $z_j(t^*) = \tilde{z}_j(t^*)$.

Therefore, combining (A.9) and (A.10), we have $z_j(t^*) \le \bar{z}_j \le z_j(t^*) + \Lambda_j(T) - \Lambda_j(t^*)$. In other words, the decision maker would still be able to achieve the hindsight optimum if she uses probabilistic allocation up to t^*, and then gets perfect information from t^* onwards, because she can accept $\bar{z}_j - z_j(t^*)$ of class j customers. If the decision maker re-solves once at t^*, then the regret can be written as

$$v^{\mathsf{HO}} - v^{\mathsf{IRT}^1} = \mathbb{E}[V^{\mathsf{HO}} - V^{\mathsf{IRT}^1}] = \mathbb{E}[V^{\mathsf{HO}} - V^{\mathsf{HO}^1}] + \mathbb{E}[V^{\mathsf{HO}^1} - V^{\mathsf{IRT}^1}]. \tag{A.11}$$

where we can decompose the revenue earned under the policy HO^1 and IRT^1 as

$$V^{\mathsf{HO}^1} = V^{\mathsf{SPA}'}(0,t^*) + V^{\mathsf{HO}}(t^*,T), \quad V^{\mathsf{IRT}^1} = V^{\mathsf{SPA}'}(0,t^*) + V^{\mathsf{SPA}}(t^*,T).$$

Consequently, we get

$$\mathbb{E}[V^{\mathsf{HO}^1} - V^{\mathsf{IRT}^1}] = \mathbb{E}[V^{\mathsf{HO}}(t^*,T) - V^{\mathsf{SPA}}(t^*,T)] = O(\sqrt{T-t^*}) = O(T^{5/12}). \tag{A.12}$$

Eq (A.12) follows the well-known result that static probabilistic allocation without re-solving has a regret rate of $O(\sqrt{k})$ for a problem with horizon length k, where the constant factor does not depend on the capacity vector C (see e.g. Reiman and Wang, 2008). For completeness, we give a proof of this result in Appendix §A.2.

Recall that if the event E happens, the hindsight optimal is still attainable starting from t^*. In other words, conditioned on E, the regret of HO^1 is $V^{\mathsf{HO}} - V^{\mathsf{HO}^1} = 0$. Therefore, the first term of (A.11) is given by

$$\begin{aligned}
\mathbb{E}[V^{\mathsf{HO}} - V^{\mathsf{HO}^1}] &= \mathbb{E}[V^{\mathsf{HO}} - V^{\mathsf{HO}^1} \mid E]\,\mathbb{P}(E) + \mathbb{E}[V^{\mathsf{HO}} - V^{\mathsf{HO}^1} \mid E^c]\,\mathbb{P}(E^c) \\
&= \mathbb{E}[V^{\mathsf{HO}} - V^{\mathsf{HO}^1} \mid E^c]\,\mathbb{P}(E^c) \\
&\le \mathbb{E}[V^{\mathsf{HO}} \mid E^c]\,\mathbb{P}(E^c).
\end{aligned} \tag{A.13}$$

Note that the hindsight optimum is bounded almost surely by $V^{\mathrm{HO}} \leq \sum_{j=1}^{n} r_j \Lambda_j(T)$, where $\Lambda_j(T)$ is the total number of arrivals from class j. Moreover, $\Lambda_j(T)$ follows Poisson distribution with mean $\lambda_j T$. By the Poisson tail bound (see Lemma 8 in Appendix A.4), we have

$$\begin{aligned}
&\mathbb{E}[(\Lambda_j(T) - 2\lambda_j T)^+] \\
&\quad = \int_0^\infty \mathbb{P}(\Lambda_j(T) - 2\lambda_j T \geq x) dx \leq \int_0^\infty 2\exp\left(-\frac{(x+\lambda_j T)^2}{2(x+2\lambda_j T)}\right) dx \\
&\quad = \int_0^{\lambda_j T} 2\exp\left(-\frac{(x+\lambda_j T)^2}{2(x+2\lambda_j T)}\right) dx + \int_{\lambda_j T}^\infty 2\exp\left(-\frac{(x+\lambda_j T)^2}{2(x+2\lambda_j T)}\right) dx \\
&\quad \leq \int_0^{\lambda_j T} 2\exp\left(-\frac{(x+\lambda_j T)^2}{2(\lambda_j T+2\lambda_j T)}\right) dx + \int_{\lambda_j T}^\infty 2\exp\left(-\frac{(x+\lambda_j T)^2}{2(x+2x)}\right) dx \\
&\quad \leq \int_0^{\lambda_j T} 2\exp\left(-\frac{(x+\lambda_j T)\lambda_j T}{6\lambda_j T}\right) dx + \int_{\lambda_j T}^\infty 2\exp\left(-\frac{(x+\lambda_j T)x}{6x}\right) dx \\
&\quad \leq \int_0^\infty 2\exp\left(-\frac{(x+\lambda_j T)}{6}\right) dx = 12\exp\left(-\frac{\lambda_j T}{6}\right).
\end{aligned}$$

Combining the above inequality with Eq (A.13), we have

$$\begin{aligned}
\mathbb{E}[V^{\mathsf{HO}} - V^{\mathsf{HO}^1}] &\leq \mathbb{E}[V^{\mathsf{HO}} \mid E^c]\,\mathbb{P}(E^c) \\
&\leq \mathbb{E}\left[\sum_{j=1}^n r_j \Lambda_j(T) \middle| E^c\right] \mathbb{P}(E^c) \\
&\leq \mathbb{E}\left[\sum_{j=1}^n r_j \left(2\lambda_j T + (\Lambda_j(T) - 2\lambda_j T)^+\right) \middle| E^c\right] \mathbb{P}(E^c) \\
&\leq \sum_{j=1}^n r_j \left(2\lambda_j T\,\mathbb{P}(E^c) + \mathbb{E}[(\Lambda_j(T) - 2\lambda_j T)^+ | E^c]\,\mathbb{P}(E^c)\right) \\
&\leq \sum_{j=1}^n r_j \left(2\lambda_j T\,\mathbb{P}(E^c) + \mathbb{E}[(\Lambda_j(T) - 2\lambda_j T)^+]\right) \\
&\leq \sum_{j=1}^n r_j \left(2\lambda_j T\,\mathbb{P}(E^c) + 12 e^{-\lambda_j T/6}\right).
\end{aligned} \tag{A.14}$$

Using the result from Lemma 10 in Appendix A.4 and (A.13)–(A.14), we have

$$\mathbb{E}[V^{\mathsf{HO}} - V^{\mathsf{HO}^1}] \leq \sum_{j=1}^{n} r_j \left(2\lambda_j T \cdot O(e^{-\kappa T^{1/6}}) + 12e^{-\lambda_j T/6}\right) = O(Te^{-\kappa T^{1/6}}). \tag{A.15}$$

We can conclude from (A.11), (A.12) and (A.15) that

$$v^{\mathsf{HO}} - v^{\mathsf{IRT}^1} = O(Te^{-\kappa T^{1/6}}) + O(T^{5/12}),$$

where the big O notation hides constants that are independent of the time horizon T and the capacity vector C. □

A.2 Additional Results

A.2.1 A note on the DLP upper bound

We establish the revenue loss of heuristics by comparing their revenues to the hindsight optimum upper bound v^{HO}. This bound is tighter than the DLP upper bound v^{DLP}. The following result suggests that v^{DLP} is not an appropriate benchmark to prove $O(1)$ revenue loss, because even the gap between the optimal policy v^* and v^{DLP} is $\Omega(\sqrt{T})$.

Proposition 10. *The gap between the optimal value of the DLP and the optimal value obtained by dynamic programming is bounded below by*

$$v^{\mathrm{DLP}} - v^* = \Omega(\sqrt{T}).$$

Proof of Proposition 10. To prove Proposition 10, we consider the following instance. In this instance, there is only one class of customer and one type of resource. So for simplicity, we will suppress the subscriptions. Suppose the expected number of arrivals in one period is Poisson process with rate λ. The revenue earned by accepting a customer is 1. The resource has the capacity λT and the amount of the resource used to serve one customer is

1. Therefore, the DLP formulation is given by

$$\max_x \left\{ Tx \,\middle|\, x \le \lambda T/T = \lambda,\ 0 \le x \le \lambda \right\}.$$

It easily verified that, we have $x^* = \lambda$ and thus $v^{\mathrm{DLP}} = \lambda T$.

On the other hand, it is obvious that the optimal policy is to admit all customers in $[0,T]$ subject to the capacity constraint. Specifically, the optimal number of the admitted customers is either the number of the arriving customers in $[0,T]$ or the capacity level, whichever is lower. Therefore, the optimal revenue of the above problem instance is given by

$$\begin{aligned} v^* &= \mathbb{E}[\min(\Lambda(T), \lambda T)] = \lambda T - \mathbb{E}[\max(\lambda T - \Lambda(T), 0)] \\ &= \lambda T - \sqrt{\lambda T}\, \mathbb{E}\left[\max\left(\frac{\lambda T - \Lambda(T)}{\sqrt{\lambda T}}, 0\right)\right]. \end{aligned}$$

From the Markov's inequality, it follows that

$$\mathbb{E}\left[\max\left(\frac{\lambda T - \Lambda(T)}{\sqrt{\lambda T}}, 0\right)\right] \ge \mathbb{P}\left(\max\left(\frac{\lambda T - \Lambda(T)}{\sqrt{\lambda T}}, 0\right) \ge 1\right).$$

Consequently, we can write

$$\begin{aligned} v^* &\le \lambda T - \sqrt{\lambda T}\, \mathbb{P}\left(\max\left(\frac{\lambda T - \Lambda(T)}{\sqrt{\lambda T}}, 0\right) \ge 1\right) \\ &= \lambda T - \sqrt{\lambda T}\, \mathbb{P}\left(\frac{\lambda T - \Lambda(T)}{\sqrt{\lambda T}} \ge 1\right) \\ &= \lambda T - \sqrt{\lambda T}(1 - F_T(1)), \end{aligned} \tag{A.16}$$

where F_T is the cumulative distribution function of $\frac{\lambda T - \Lambda(T)}{\sqrt{\lambda T}}$. Let Φ denote the cumulative distribution function of standard normal distribution. Since

$$1 - F_T(1) = 1 - \Phi(1) + \Phi(1) - F_T(1) \geq 1 - \Phi(1) - |F_T(1) - \Phi(1)|. \tag{A.17}$$

We use the Berry-Esseen theorem (Lemma 4 in Appendix A.4) to bound $|F_T(1) - \Phi(1)|$. Let $X_i = \lambda - (\Lambda(i) - \Lambda(i-1))$ for $i = 1, \ldots, T$. From the stationary and independent increment properties of Poisson processes, we observe that X_i are i.i.d. with $\mathbb{E}[X_1] = 0$, $\mathbb{E}[X^2] = \lambda$ and $\mathbb{E}[|X_1^3|] = \lambda$. Since

$$\frac{X_1 + \ldots + X_T}{\sqrt{\lambda T}} = \frac{\lambda T - \Lambda(T)}{\sqrt{\lambda T}},$$

from Lemma 4 we have

$$|F_T(1) - \Phi(1)| \leq \frac{0.4748\lambda}{\sqrt{\lambda^3 T}} = \frac{0.4748}{\sqrt{\lambda T}}. \tag{A.18}$$

Combining (A.16), (A.17) and (A.18), we can write

$$v^* \leq \lambda T - \sqrt{\lambda T}\left(1 - \Phi(1) - \frac{0.4748}{\sqrt{\lambda T}}\right) = \lambda T - \sqrt{\lambda T}(1 - \Phi(1)) + 0.4748.$$

Since $v^{\mathrm{DLP}} = \lambda T$ and $1 - \Phi(1) = 0.1587$, it follows that

$$v^{\mathrm{DLP}} - v^* \geq 0.1587\sqrt{\lambda T} - 0.4748 = \Omega(\sqrt{T}).$$

□

A.2.2 Revenue loss of static probabilistic allocation

The following result is a well-known in the revenue management literature by Gallego and van Ryzin (1994) and Gallego and Ryzin (1997); see also Cooper (2002) and Reiman and Wang (2008).

Proposition 11. *The gap between the optimal value of the DLP and the optimal value* *obtained by the static probabilistic allocation (*SPA*) heuristic is bounded above by*

$$v^{\mathrm{DLP}} - v^{\mathsf{SPA}} = O(\sqrt{T}).$$

The constant pre-factor depends on the customer arrival rate λ_j ($\forall j \in [n]$), the revenues per customer r_j ($\forall j \in [n]$), and the BOM matrix A; however, it does not depend on the starting capacity C_l ($\forall l \in [m]$).

Since the revenue of the optimal policy, v^*, and the hindsight optimum, v^{HO}, satisfies $v^* \leq v^{\mathrm{HO}} \leq v^{\mathrm{DLP}}$, a corollary of the result is $v^* - v^{\mathsf{SPA}} = O(\sqrt{T})$ and $v^{\mathrm{HO}} - v^{\mathsf{SPA}} = O(\sqrt{T})$.

In the revenue management literature, this result is often proved under the additional assumption that resource capacities and customer arrivals are both scaled up at the same rate. However, the result in fact holds for arbitrary capacity levels. We need this fact in the proof of Theorem 1. To make the proof of Theorem 1 self-contained, we include a proof of the proposition below.

Proof of Proposition 11. By Eq 2.2, we have $v^{\mathrm{DLP}} = T\sum_{j\in[n]} r_j x_j^*$. We bound v^{SPA} in two steps. First, consider a hypothetical setting where remaining capacities are allowed to become negative. Since SPA accepts each class j customer with probability x_j^*/λ_j if capacity constraints are ignored, the expected revenue of SPA under this hypothetical setting is

$$\mathbb{E}\left[\sum_{j\in[n]} \int_0^T r_j \lambda_j \frac{x_j^*}{\lambda_j} dt\right] = \sum_{j\in[n]} r_j x_j^* T = v^{\mathrm{DLP}}.$$

In reality, remaining capacity is always nonnegative, and customers must be rejected if there is insufficient capacity. So to correct the revenue calculation of the hypothetical setting, we must subtract the revenue associated with customers who are rejected due to insufficient capacity. This part of revenue is bounded by

$$\sum_{l=1}^{m} r_{max}^l \mathbb{E}\big[\big(\sum_{j=1}^{n} a_{lj} X_j - C_l\big)^+\big], \tag{A.19}$$

where X_j is the number of class j customers that would have been accepted by SPA without capacity limits, which follow a Poisson distribution with mean $x_j^* T$, and r_{max}^l is the largest possible revenue gain by increasing the capacity of resource l by one unit, i.e., $r_{max}^l = \max_{j\in[n]}\{r_j \mathrm{I}(a_{lj} > 0)/a_{lj}\}$. We have

$$\begin{aligned}\mathbb{E}\big[\big(\sum_{j=1}^{n} a_{lj}X_j - C_l\big)^+\big] &\le \mathbb{E}\big[\big(\sum_{j=1}^{n} a_{lj}X_j - \sum_{j=1}^{n} a_{lj}x_j^* T\big)^+\big] \\ &\le \mathbb{E}\big[\big|\sum_{j=1}^{n} a_{lj}X_j - \sum_{j=1}^{n} a_{lj}x_j^* T\big|\big] \\ &\le \sqrt{\mathbb{E}\big[\big(\sum_{j=1}^{n} a_{lj}X_j - \sum_{j=1}^{n} a_{lj}x_j^* T\big)^2\big]},\end{aligned}$$

where the first inequality follows the capacity constraints in DLP (2.2), and the last inequality follows from Cauchy-Schwarz inequality.

Note that X_j's mean and variance are both equal to $x_j^* T$, and X_j's are independent for all $j \in [n]$. So

$$\begin{aligned}\mathbb{E}\big[\big(\sum_{j=1}^{n} a_{lj}X_j - \sum_{j=1}^{n} a_{lj}x_j^* T\big)^2\big] &= \mathrm{Var}\big(\sum_{j=1}^{n} a_{lj}X_j\big) = \sum_{j=1}^{n} a_{lj}^2 \mathrm{Var}\big(X_j\big) \\ &= \sum_{j=1}^{n} a_{lj}^2 x_j^* T \le \sum_{j=1}^{n} a_{lj}^2 \lambda_j T,\end{aligned}$$

where the last inequality follows the demand constraints $x_j^* \le \lambda_j$ in DLP (2.2). Substituting this result to Eq (A.19), we have

$$v^{\mathsf{DLP}} - v^{\mathsf{SPA}} \le \sum_{l=1}^{m} r_{max}^l \mathbb{E}\big[\big(\sum_{j=1}^{n} a_{lj}X_j - C_l\big)^+\big] \le \sum_{l=1}^{m} r_{max}^l \sqrt{\sum_{j=1}^{n} a_{lj}^2 \lambda_j T}.$$

□

A.3 Proofs for Results in Section 2.5

In this section, we provide complete proofs for the results on the FR policy in Section 2.5.

A.3.1 Proof of Proposition 2

Proof of Proposition 2. Let V^{FR} denote the revenue earned under the FR policy. We know that $v^{\mathrm{HO}} - v^{\mathsf{FR}} = \mathbb{E}[V^{\mathrm{HO}}] - \mathbb{E}[V^{\mathsf{FR}}] = \mathbb{E}[V^{\mathrm{HO}} - V^{\mathsf{FR}}]$. From the law of total expectation, it equals to

$$\mathbb{E}[V^{\mathrm{HO}} - V^{\mathsf{FR}}|Q]\,\mathbb{P}(Q) + \mathbb{E}[V^{\mathrm{HO}} - V^{\mathsf{FR}}|Q^c]\,\mathbb{P}(Q^c), \tag{A.20}$$

for any event Q. Since V^{HO} is an upper bound of V^{FR}, i.e., $V^{\mathrm{HO}} \geq V^{\mathsf{FR}}$ a.s. and the probability of any measurable event is nonnegative, the second term of (A.20) is nonnegative. Consequently,

$$v^{\mathrm{HO}} - v^{\mathsf{FR}} \geq \mathbb{E}[V^{\mathrm{HO}} - V^{\mathsf{FR}}|Q]\,\mathbb{P}(Q).$$

That is, to complete the proof, we want to show that $\mathbb{E}[V^{\mathrm{HO}} - V^{\mathsf{FR}}|Q]\,\mathbb{P}(Q) \geq \Omega(\sqrt{T})$ for some appropriately chosen event Q.

Recall that we consider a problem instance with two classes of customers and one resource. Customers from each class arrive according to a Poisson process with rate 1. The arrivals from two classes are assumed to be independent. The initial resource capacity is T. Customers from both classes, if accepted, consume one unit of the resource, but pay different prices, $r_1 > r_2$. To proof the result, we will consider the situation when the number of arrivals of class 1 customers in $[0,T]$ is above its mean which is T. Since the initial capacity of the resource is T, the hindsight optimal policy is to accept only class 1 customer. More specifically, we should accept T of class 1 customer and none of class 2 customer. Failing to do so will result in a positive regret.

We will partition time in the interval $[0,T]$ into 3 phases of equal length. Let T' and T'' denote the beginning of phase 2 and phase 3 respectively. In other words, $T' = T/3$ and $T'' = 2T/3$. We will define the events of the number of the arrivals of class 1 customer in each period as follows.

$$Q_1 = \{T' - 4\sqrt{T'} \leq \Lambda_1(0,T') \leq T' - 3\sqrt{T'}\}, \tag{A.21}$$

$$Q_2 = \{(t-T') - 2\sqrt{T'} \leq \Lambda_1(T',t) \leq (t-T') + 2\sqrt{T'},\ \forall t \in (T',T'']\}, \tag{A.22}$$

$$Q_3 = \{T' + 6\sqrt{T'} \leq \Lambda_1(T'',T) \leq T' + 7\sqrt{T'}\}, \tag{A.23}$$

where Q_i restricts the number of the arrivals of class 1 customer in phase i. Let $z_j(t_1,t_2)$ denote the actual number of class j customers admitted in $(t_1,t_2]$. We will further define the events of the number of accepted customers as

$$B = \{z_2(0,T) \geq \frac{1}{6}\sqrt{T'}\}, \tag{A.24}$$

$$B_1 = \{z_1(0,T') < T' - 10\sqrt{T'}\}. \tag{A.25}$$

If the event B happens, the decision maker will admit at least $\frac{1}{6}\sqrt{T'}$ of class 2 customers which leads to a regret of at least $(r_1 - r_2)\frac{1}{6}\sqrt{T'}$ from the hindsight optimal. On the other hand, if the event B_1 happens, the decision maker will admit less than $T' - 10\sqrt{T'}$ of class 1 customer; this means that even if the decision maker admit all arrivals of class 1 customer in the second and the third phase, the total number of admitted class 1 customer is

$$z_1(0,T) < T' - 10\sqrt{T'} + T' + 2\sqrt{T'} + T' + 7\sqrt{T'} = T - \sqrt{T'},$$

which results in a regret of at least $(r_1 - r_2)\sqrt{T'}$ from the hindsight optimal. Therefore, if the event B or B_1 happens with probability that is bounded away from zero, the incurred a regret of is at least $(r_1 - r_2)\frac{1}{6}\sqrt{T'}$ from the hindsight optimal. If we can show that this event happens with positive probability, then we are done. That is, we want to show that

$$\mathbb{P}((B \cup B_1) \cap Q_1 \cap Q_2 \cap Q_3) > 0.$$

This probability can be written as

$$
\begin{aligned}
\mathbb{P}((B\cup B_1)\cap Q_1\cap Q_2\cap Q_3) &= \mathbb{P}(Q_1\cap Q_2\cap Q_3)-\mathbb{P}(B^c\cap B_1^c\cap Q_1\cap Q_2\cap Q_3) \\
&= \mathbb{P}(Q_1)\,\mathbb{P}(Q_2)\,\mathbb{P}(Q_3)-\mathbb{P}(B^c\cap B_1^c\cap Q_1\cap Q_2\cap Q_3) \quad \text{(A.26)}\\
&\geq \mathbb{P}(Q_1)\,\mathbb{P}(Q_2)\,\mathbb{P}(Q_3)-\mathbb{P}(B^c\cap B_1^c\cap Q_1\cap Q_2), \quad \text{(A.27)}
\end{aligned}
$$

where the first term of (A.26) follows since the arrivals of the customers in disjoint interval are independent, i.e., Q_1, Q_2 and Q_3 are independent, and the inequality (A.27) follows because the event $B^c\cap B_1^c\cap Q_1\cap Q_2\cap Q_3$ is a subset of the event $B^c\cap B_1^c\cap Q_1\cap Q_2$. The remaining part of the proof relies on the results of Lemma 11 in Appendix A.4. Combining the result from Lemma 11 and (A.27), we get

$$
\begin{aligned}
&\mathbb{P}((B\cup B_1)\cap Q_1\cap Q_2\cap Q_3)\\
&\qquad\geq \left(0.0013-\frac{0.9496}{\sqrt{T'}}\right)(0.5)\left(9.8531\times 10^{-10}-\frac{0.9496}{\sqrt{T'}}\right)-e^{-0.0026\sqrt{T'}}.
\end{aligned}
$$

Therefore, the regret of FR is given by

$$
\begin{aligned}
&v^{\mathsf{HO}}-v^{\mathsf{FR}}\\
&\quad\geq (r_1-r_2)\frac{1}{6}\sqrt{T'}\\
&\qquad\cdot\left(\left(0.0013-\frac{0.9496}{\sqrt{T'}}\right)(0.5)\left(9.8531\times 10^{-10}-\frac{0.9496}{\sqrt{T'}}\right)-e^{-0.0026\sqrt{T'}}\right)\\
&\quad=\Omega(\sqrt{T}).
\end{aligned}
$$

□

A.3.2 Proof of Proposition 3

Proof Proposition 3. The FR algorithm divides the horizon $[0,T]$ into T periods: $[0,1)\cup[1,2)\cup\cdots\cup[T-1,T]$. In each period $[t,t+1)$, the algorithm attempts to accept class j

customers with probability $x_j(t)/\lambda_j$. If we ignore capacity constraints, the algorithm on average accepts $x_j(t)$ customers from class j. However, the decision maker can potentially reject customers due to capacity constraints; if that happens, the expected number of admitted class j customers in period t is less than $x_j(t)$ per period.

More specifically, let us consider the LP solved at time t:

$$\max_y \left\{ \sum_{j=1}^n r_j y_j \,\middle|\, \sum_{j=1}^n A_j y_j \le C(t), \text{ and } 0 \le y_j \le \lambda_j(T-t), \forall j \in [n] \right\}.$$

Let $y(t)$ be the optimal solution. The algorithm accepts customers in class j with probability $\frac{x_j(t)}{\lambda_j} = \frac{y_j(t)}{\lambda_j(T-t)}$. If there is insufficient capacity for class j in period $[t,t+1)$, two cases can happen: case 1) $A_j \nleq C(t)$, i.e., there exists $l \in [m]$ such that $a_{lj} > C_l(t)$, so there is insufficient capacity for class j when this period starts; case 2) $A_j \le C(t)$, namely there is sufficient capacity when this period starts, but during $[t,t+1)$ the capacity of certain resource runs out, and a class j customer that arrives at time $s \in [t,t+1)$ finds $A_j \nleq C(s)$. For case 1), $A_j \nleq C(t)$ implies $y_j(t) < 1$. So, the expected number of class j customers that FR would accept, but that are not accepted because of the capacity constraint in period $[t,t+1)$ is less than $\lambda_j \cdot \frac{1}{\lambda_j(T-t)} = \frac{1}{T-t}$. The revenue loss from that group of customers over the entire horizon is bounded by

$$\sum_{j=1}^n \sum_{t=0}^{T-1} r_j \frac{1}{T-t} \le \sum_{j=1}^n r_j (\log T + 1),$$

where we use the fact that $\sum_{i=2}^T \frac{1}{i} \le \int_{x=1}^T \frac{1}{x} dx = \log(T)$. For case 2), we note that this situation can happen *at most once* during the entire horizon for each class of customers. Since the expected number of class j customers that FR would accept, but that are not accepted because of the capacity constraint in one period is bounded above by the expected number of class j arrivals in one period which is λ_j, the revenue loss caused by case 2) is bounded by $\sum_{j=1}^n r_j \lambda_j$.

We can write the expected revenue of the re-solving heuristic as

$$v^{\mathsf{FR}} \geq \mathbb{E}\left[\sum_{t=0}^{T-1}\sum_{j=1}^{n} r_j x_j(t)\right] - \sum_{j=1}^{n} r_j(\log T + 1) - \sum_{j=1}^{n} r_j \lambda_j,$$

where the last two terms account for the lost sales of case 1) and case 2) respectively. Therefore, the expected revenue loss of the re-solving heuristic can be bounded by

$$v^{\mathsf{DLP}} - v^{\mathsf{FR}} \leq T\sum_{j=1}^{n} r_j x_j^* - \mathbb{E}\left[\sum_{t=0}^{T-1}\sum_{j=1}^{n} r_j x_j(t)\right] + \sum_{j=1}^{n} r_j(\log T + 1) + \sum_{j=1}^{n} r_j \lambda_j. \tag{A.28}$$

Since the solutions to the DLP(x^*) and the LP solved under the FR policy($x(t)$) only differ in the right hand side of the capacity constraints (b and $b(t)$), we can write, for any time period $[t, t+1)$,

$$\sum_{j=1}^{n} r_j x_j^* - \sum_{j=1}^{n} r_j x_j(t) \leq \sum_{l=1}^{L} r_{max}^l (b_l - b_l(t))^+, \tag{A.29}$$

where r_{max}^l is the largest possible revenue gain by increasing the capacity of resource l by one unit, i.e., $r_{max}^l = \max_{j\in[n]}\{r_j \mathrm{I}(a_{lj} > 0)/a_{lj}\}$. Equation (A.29) holds because r_{max}^l is an upper bound of the dual price for resource $l \in [m]$. From the definition of $b_l(i)$, we can write

$$(b_l - b_l(t))^+ = \left[\sum_{i=0}^{t-1}(b_l(i) - b_l(i+1))\right]^+ = \left[\sum_{i=0}^{t-1}\left(\frac{C_l(i)}{T-i} - \frac{C_l(i+1)}{T-i-1}\right)\right]^+.$$

Note that $\frac{1}{T-i} = \frac{1}{T-i-1} - \frac{1}{(T-i-1)(T-i)}$, so it follows that

$$\begin{aligned}(b_l - b_l(t))^+ &= \left[\sum_{i=0}^{t-1}\left(\frac{C_l(i)}{T-i-1} - \frac{C_l(i)}{(T-i-1)(T-i)} - \frac{C_l(i+1)}{T-i-1}\right)\right]^+ \\ &= \left[\sum_{i=0}^{t-1}\left(\frac{C_l(i) - C_l(i+1)}{T-i-1} - \frac{C_l(i)}{(T-i-1)(T-i)}\right)\right]^+.\end{aligned} \tag{A.30}$$

Let $z_j(t)$ be the actual number of class j customers admitted in $[0, t]$. The change in the capacity of resource l is given by $C_l(i) - C_l(i+1) = \sum_{j=1}^{n} a_{lj}(z_j(i+1) - z_j(i))$. Because of the capacity constraint, the decision maker may fail to accept some customers. Therefore, the actual number of the admitted customers in any period is bounded above by the number of the customers admitted in that period by ignoring the capacity constraint.

More specifically, we define stochastic processes $\{\tilde{z}_j(t), t \geq 1, j \in [n]\}$, such that $\tilde{z}_j(t+1) - \tilde{z}_j(t)$ follows Poisson distribution with mean $x_j(t)$ (the solution to the LP at period t). Therefore, $\tilde{z}_j(t+1) - \tilde{z}_j(t)$ is the the number of class j customers that the algorithm could have admitted if there were no capacity constraint in $[t, t+1)$. Since the number of customers who are actually admitted, $z_j(t+1) - z_j(t)$, follows the same Poisson distribution with additional rejections due to capacity constraints, we always have $z_j(t+1) - z_j(t) \leq \tilde{z}_j(t+1) - \tilde{z}_j(t)$, and therefore

$$C_l(i) - C_l(i+1) = \sum_{j=1}^{n} a_{lj}(z_j(i+1) - z_j(i)) \leq \sum_{j=1}^{n} a_{lj}(\tilde{z}_j(i+1) - \tilde{z}_j(i)).$$

We can now bound (A.30) by

$$(b_l - b_l(t))^+ \leq \left[\sum_{i=0}^{t-1} \left(\frac{\sum_{j=1}^{n} a_{lj}(\tilde{z}_j(i+1) - \tilde{z}_j(i))}{T-i-1} - \frac{b_l(i)}{T-i-1} \right) \right]^+ \tag{A.31}$$

$$\leq \left[\sum_{i=0}^{t-1} \left(\frac{\sum_{j=1}^{n} a_{lj}(\tilde{z}_j(i+1) - \tilde{z}_j(i))}{T-i-1} - \frac{\sum_{j=1}^{n} a_{lj} x_j(i)}{T-i-1} \right) \right]^+, \tag{A.32}$$

where the second term in the RHS of (A.31) follows from the definition of $b_l(i) = \frac{C_l(i)}{T-i}$ and the inequality of (A.32) follows from the definition of the re-solving LP in Algorithm 2 which is $\sum_{j=1}^{n} a_{lj} x_j(i) \leq b_l(i)$.

We will use the result from Lemma 12 in Appendix A.4 to finish the proof. If we sum the inequality (A.29) over $t \in \{0\} \cup [T-1]$ and apply Lemma 12, it follows that

$$
\begin{aligned}
& T\sum_{j=1}^{n} r_j x_j^* - \mathbb{E}\left[\sum_{t=0}^{T-1}\sum_{j=1}^{n} r_j x_j(t)\right] \\
& \le \mathbb{E}\left[\sum_{t=0}^{T-1}\sum_{l=1}^{L} r_{max}^l (b_l - b_l(t))^+\right] \\
& \le \sum_{l=1}^{L} r_{max}^l K_l \sum_{t=0}^{T-1}\sqrt{\sum_{i=0}^{t-1}\frac{1}{(T-i-1)^2}} \\
& = \sum_{l=1}^{L} r_{max}^l K_l \sum_{t=1}^{T}\sqrt{\sum_{i=1}^{t-1}\frac{1}{(T-i)^2}} \\
& = \sum_{l=1}^{L} r_{max}^l K_l \left[\sum_{t=1}^{T-1}\sqrt{\sum_{i=1}^{t-1}\frac{1}{(T-i)^2}} + \sqrt{\sum_{i=1}^{T-1}\frac{1}{(T-i)^2}}\right] \\
& \le \sum_{l=1}^{L} r_{max}^l K_l \left[\sum_{t=1}^{T-1}\sqrt{\int_1^t \frac{1}{(T-s)^2}ds} + \sqrt{\int_1^{T-1}\frac{1}{(T-s)^2}ds + 1}\right] \\
& \le \sum_{l=1}^{L} r_{max}^l K_l \left[\sum_{t=1}^{T-1}\sqrt{\frac{1}{T-t}} + \sqrt{1 - \frac{1}{T-1} + 1}\right] \\
& \le \sum_{l=1}^{L} r_{max}^l K_l (2\sqrt{T} + \sqrt{2}).
\end{aligned}
$$

Combining the result to (A.28), we can conclude that

$$
v^{\mathsf{DLP}} - v^{\mathsf{FR}} \le \sum_{l=1}^{L} r_{max}^l K_l (2\sqrt{T} + \sqrt{2}) + n r_{max}(\log T + 1) + n r_{max}\lambda_{max} = O(\sqrt{T}),
$$

where $r_{max} = \max_{j\in[n]} r_j$ and $\lambda_{max} = \max_{j\in[n]} \lambda_j$. $\square$

A.4 Lemmas

Lemma 4 (Berry-Esseen theorem, Corollary 1 in Shevtsova (2011))**.** *Let $X_1, X_2, \ldots$ be independent and identically distributed random variables (i.i.d.) with $\mathbb{E}[X_1] = 0$, $\mathbb{E}[X_1^2] = \sigma^2 > 0$ and $\mathbb{E}[|X_1^3|] = \rho < \infty$. Let F_n be the cumulative distribution function of $\frac{X_1+\ldots+X_n}{\sigma\sqrt{n}}$ and Φ the cumulative distribution function of standard normal distribution. For any x and* *n*,

$$|F_n(x) - \Phi(x)| \leq \frac{0.4748\rho}{\sigma^3\sqrt{n}}.$$

Lemma 5 (Doob's maximal inequality, Theorem 1.7 in Revuz and Yor (1999), p. 54)**.** *Suppose M_t is a martingale with paths that are right continuous with left limits. Then, for any constant $a > 0$,*

$$\mathbb{P}(\sup_{s \leq t} |M_s| \geq a) \leq \frac{\mathbb{E}[|M_t|]}{a}.$$

Lemma 6 (Freedman (1975), Theorem 4.1)**.** *Given a sequence of real-valued supermartingale differences $(\xi_i, \mathscr{F}_i)_{i \in \{0\} \cup [n]}$ with $\xi_0 = 0$. Set $S_k = \sum_{i=0}^{k} \xi_i$ for $k \in [n]$. Then,*

$$S = (S_k, \mathscr{F}_k)_{k \in [n]} \text{ is a supermartingale.}$$

Let $\langle S \rangle_k = \sum_{i=1}^{k} \mathbb{E}[\xi_i^2 | \mathscr{F}_{i-1}]$. Suppose $\xi_i \leq 1$. Then, for all $x, v > 0$,

$$\mathbb{P}(\exists k \in [n] : S_k \geq x, \langle S \rangle_k \leq v^2) \leq \exp\left(-\frac{x^2}{2(v^2 + x)}\right).$$

Lemma 7 (Chernoff Bound for Poisson Distribution, Boucheron, Lugosi, and Massart (2013), p. 22–23)**.** *Suppose random variable X follows a Poisson distribution with parameter $\lambda > 0$, for any constant $x > 0$, we have*

$$\mathbb{P}(X \geq \lambda + x) \leq \exp\left(-\frac{x^2}{2\lambda} h\left(\frac{x}{\lambda}\right)\right), \tag{A.33}$$

and, for any constant $0 < x < \lambda$,

$$\mathbb{P}(X \leq \lambda - x) \leq \exp\left(-\frac{x^2}{2\lambda} h\left(-\frac{x}{\lambda}\right)\right), \tag{A.34}$$

where $h(t) = 2((1+t)\log(1+t) - t)/t^2$ for any $t > -1$.

Lemma 8 (Two-sided Poisson Tail Bound)**.** *If random variable X follows a Poisson distri-*

bution with parameter $\lambda > 0$*, for any constant* $x > 0$*, we have*

$$\mathbb{P}(|X-\lambda| \geq x) \leq 2\exp\left(-\frac{x^2}{2(\lambda+x)}\right).$$

Proof of Lemma 8. This result is well-known and directly follows from Lemma 7. For completeness, we provide a proof here.

First, we want to show that, for $t > 0$, $\frac{t^2}{2}h(t) \geq \frac{t^2}{2(1+t)}$, where $h(t) = 2\frac{(1+t)\log(1+t)-t}{t^2}$. To this end, we define $g(t) = (1+t)t^2h(t) - t^2 = 2(1+t)^2\log(1+t) - 2t - 3t^2$, and prove that $g(t) \geq 0$. We have $g(0) = 0$, $g'(t) = 4(1+t)\log(1+t) - 4t$, and $g''(t) = 4\log(1+t)$. Therefore, $g(t)$ is non-decreasing and convex for $t > 0$. We can conclude that $g(t) \geq 0$, and hence, $\frac{t^2}{2}h(t) \geq \frac{t^2}{2(1+t)}$. Using this fact, we have, for all $x > 0$, $\frac{x^2}{2\lambda}h\left(\frac{x}{\lambda}\right) \geq \frac{x^2}{2(\lambda+x)}$. Therefore, from (A.33), we get $\mathbb{P}(X \geq \lambda + x) \leq \exp\left(-\frac{x^2}{2\lambda}h\left(\frac{x}{\lambda}\right)\right) \leq \exp\left(-\frac{x^2}{2(\lambda+x)}\right)$.

Next, we want to show that for $-1 < t < 0$, it holds that $\frac{t^2}{2}h(t) \geq \frac{t^2}{2(1-t)}$. To this end, we define $f(t) = (1-t)t^2h(t) - t^2 = 2(1-t^2)\log(1+t) - 2t + t^2$, and prove that $f(t) \geq 0$ for $-1 < t < 0$. We have $f(0) = 0$ and $f'(t) = -4t\log(1+t) < 0$ for $-1 < t < 0$. We can therefore conclude that $f(t) \geq 0$ for $-1 < t < 0$, and hence $\frac{t^2}{2}h(t) \geq \frac{t^2}{2(1-t)}$. Combining with (A.34), for $0 < x < \lambda$, we get $\mathbb{P}(X \leq \lambda - x) \leq \exp\left(-\frac{x^2}{2\lambda}h\left(-\frac{x}{\lambda}\right)\right) \leq \exp\left(-\frac{x^2}{2(\lambda+x)}\right)$. Combining these two cases, we obtain the result. □

Lemma 9. *There exists an optimal solution to the hindsight LP defined in* (2.3)*,* $\bar{z} = (\bar{z}_1, \ldots, \bar{z}_n)$*, such that*

$$\bar{z}_j \in [\max\{Tx_j^* - \Gamma(T), 0\}, \min\{Tx_j^* + \Gamma(T), \Lambda_j(T)\}] \quad \textit{for all } j \in [n]. \tag{A.35}$$

Proof of Lemma 9. Theorem 4.2 in Reiman and Wang (2008) shows that there exists an optimal solution to the hindsight LP defined in (2.3), $\bar{z} = (\bar{z}_1, \ldots, \bar{z}_n)$, such that

$$\bar{z}_j \in [Tx_j^* - \Gamma(T), Tx_j^* + \Gamma(T)] \quad \text{for all } j \in [n].$$

The lemma immediately follows since any feasible solution to the hindsight LP satisfies $0 \le \bar{z}_j \le \Lambda_j(T)$ for all $j \in [n]$. $\square$

Lemma 10. *The probability of event E defined in* (A.8) *satisfies $\mathbb{P}(E^c) = O\big(n|J_\lambda|\exp\big(-\kappa T^{1/6}\big)\big)$, where the constant κ is given by $\kappa = \frac{\lambda_{min}}{24(\alpha|J_\lambda|+1)^2}$. The set $J_\lambda = \{j : x_j^* = \lambda_j\}$, and α is a positive constant that depends on the BOM matrix A.*

Proof of Lemma 10. We can write

$$\begin{aligned}
\mathbb{P}(E^c) &= \mathbb{P}\Big(\bigcup_{j:x_j^* \ge \lambda_j T^{-1/4}} E_{1,j}^c \cup \bigcup_{j:x_j^* \le \lambda_j(1-T^{-1/4})} E_{2,j}^c\Big) \\
&\le \sum_{j:x_j^* \ge \lambda_j T^{-1/4}} \mathbb{P}(E_{1,j}^c) + \sum_{j:x_j^* \le \lambda_j(1-T^{-1/4})} \mathbb{P}(E_{2,j}^c). \qquad \text{(A.36)}
\end{aligned}$$

First, we will bound

$$\mathbb{P}(E_{1,j}^c) = \mathbb{P}(\tilde{z}_j(t^*) - t^* x_j^* + \Gamma(T) > (T-t)x_j^*).$$

Observe that if the event $E_{1,j}^c$ happens, at least one of the following two events must happen:

$$\left\{\tilde{z}_j(t^*) - t^* x_j^* > \frac{(T-t^*)x_j^*}{2}\right\}, \left\{\Gamma(T) > \frac{(T-t^*)x_j^*}{2}\right\}.$$

Thus, we can apply the union bound and write

$$\mathbb{P}(E_{1,j}^c) \le \mathbb{P}\left(\tilde{z}_j(t^*) - t^* x_j^* > \frac{(T-t^*)x_j^*}{2}\right) + \mathbb{P}\left(\Gamma(T) > \frac{(T-t^*)x_j^*}{2}\right). \qquad \text{(A.37)}$$

Let us consider three cases: 1) $x_j^* < \lambda_j T^{-1/4}$, 2) $\lambda_j T^{-1/4} \le x_j^* \le \lambda_j(1-T^{-1/4})$, 3) $x_j^* > \lambda_j(1-T^{-1/4})$. Case 1) is already eliminated in definition of event E, so we focus on case 2) and 3). In Case 2), $\tilde{z}_j(t^*)$ is Poisson random variable with parameter $t^* x_j^*$. We use the two-sided Poisson tail bound (Lemma 8) to bound such events.

It follows from Lemma 8 that

$$\begin{aligned}
\mathbb{P}\left(\tilde{z}_j(t^*) - t^* x_j^* > \frac{(T-t^*)x_j^*}{2}\right) &\leq 2\exp\left(-\frac{1}{2}\left(\frac{(T-t^*)x_j^*}{2}\right)^2 \frac{1}{(t^* x_j^* + (T-t^*)x_j^*/2)}\right) \\
&\leq 2\exp\left(-\frac{(T-t^*)^2 x_j^*}{8T}\right) \\
&\leq 2\exp\left(-\frac{(T^{5/6})^2 \cdot T^{-1/4}\lambda_j}{8T}\right) \\
&\leq 2\exp\left(-\frac{\lambda_{min}}{8}T^{5/12}\right), \qquad \text{(A.38)}
\end{aligned}$$

where $\lambda_{min} = \min_{j\in[n]} \lambda_j$. In case 3), $\tilde{z}_j(t^*)$ is Poisson random variable with parameter $t^*\lambda_j$. We have

$$\begin{aligned}
\mathbb{P}\left(\tilde{z}_j(t^*) - t^* x_j^* > \frac{(T-t^*)x_j^*}{2}\right) &= \mathbb{P}\left(\tilde{z}_j(t^*) - t^*\lambda_j > \frac{(T-t^*)x_j^*}{2} - t^*(\lambda_j - x_j^*)\right) \\
&\leq \mathbb{P}\left(\tilde{z}_j(t^*) - t^*\lambda_j > \frac{T^{5/6}\lambda_j/2}{2} - \lambda_j T \cdot T^{-1/4}\right).
\end{aligned}$$

Choose a constant T_0 such that $T^{5/6}/4 \geq 2T^{3/4}$ for $T \geq T_0$. It follows from Lemma 8 that

$$\begin{aligned}
\mathbb{P}\left(\tilde{z}_j(t^*) - t^*\lambda_j > \frac{T^{5/6}\lambda_j/2}{2} - \lambda_j T \cdot T^{-1/4}\right) &\leq 2\exp\left(-\frac{1}{2}\left(T^{3/4}\lambda_j\right)^2 \frac{1}{(t^*\lambda_j + T^{3/4}\lambda_j)}\right) \\
&\leq 2\exp\left(-\frac{1}{4}\frac{\lambda_j T^{3/2}}{T}\right) \\
&\leq 2\exp\left(-\frac{\lambda_{min}}{4}T^{1/2}\right). \qquad \text{(A.39)}
\end{aligned}$$

Let $J_\lambda = \{j : x_j^* = \lambda_j\}$. The second term can be written as

$$\begin{aligned}
\mathbb{P}\left(\Gamma(T) > \frac{(T-t^*)x_j^*}{2}\right) &= \mathbb{P}\left(\alpha \sum_{j\in J_\lambda} |\Lambda_j(T) - \lambda_j T| > \frac{(T-t^*)x_j^*}{2}\right) \\
&\leq \sum_{j\in J_\lambda} \mathbb{P}\left(\alpha|\Lambda_j(T) - \lambda_j T| > \frac{(T-t^*)x_j^*}{2|J_\lambda|} \cdot I(|J_\lambda| > 0)\right) \\
&= \sum_{j\in J_\lambda} \mathbb{P}\left(|\Lambda_j(T) - \lambda_j T| > \frac{(T-t^*)x_j^*}{2\alpha|J_\lambda|+1}\right).
\end{aligned}$$

Applying Lemma 8 again, we have

$$\begin{aligned}
&\mathbb{P}\left(|\Lambda_j(T) - \lambda_j T| > \frac{(T-t^*)x_j^*}{2\alpha|J_\lambda|+1}\right) \\
&2\exp\left(-\frac{1}{2}\left(\frac{(T-t^*)x_j^*}{2\alpha|J_\lambda|+1}\right)^2 \frac{1}{(\lambda_j T + (T-t^*)x_j^*/(2\alpha|J_\lambda|+1))}\right) \\
&\leq 2\exp\left(-\frac{1}{2}\frac{((T-t^*)x_j^*)^2}{2\alpha|J_\lambda|+1}\frac{1}{2(\alpha|J_\lambda|+1)\lambda_j T}\right) \\
&\leq 2\exp\left(-\frac{(T^{5/6}\cdot\lambda_j T^{-1/4})^2}{4(\alpha|J_\lambda|+1)(2\alpha|J_\lambda|+1)\lambda_j T}\right) \\
&\leq 2\exp\left(-\frac{\lambda_{min}}{4(\alpha|J_\lambda|+1)(2\alpha|J_\lambda|+1)} T^{1/6}\right).
\end{aligned}$$

Therefore, we have

$$\mathbb{P}\left(\Gamma(T) > \frac{(T-t^*)x_j^*}{2}\right) \leq 2|J_\lambda| \exp\left(-\frac{\lambda_{min}}{4(\alpha|J_\lambda|+1)(2\alpha|J_\lambda|+1)} T^{1/6}\right). \tag{A.40}$$

From (A.37), (A.38),(A.39) and (A.40), it follows that

$$\mathbb{P}(E_{1,j}^c) \leq O\left(\exp\left(-\frac{\lambda_{min}}{4(\alpha|J_\lambda|+1)(2\alpha|J_\lambda|+1)} T^{1/6}\right)\right). \tag{A.41}$$

We can also apply the similar argument to bound $\mathbb{P}(E_{2,j}^c)$. That is, if the event $E_{2,j}^c$

happens, at least one of the following three events must happen:

$$\left\{t^*x_j^* - \tilde{z}_j(t^*) > \tfrac{(T-t^*)(\lambda_j - x_j^*)}{3}\right\},\ \left\{\Gamma(T) > \tfrac{(T-t^*)(\lambda_j - x_j^*)}{3}\right\},\ \left\{|\Delta_j(t^*)| > \tfrac{(T-t^*)(\lambda_j - x_j^*)}{3}\right\}.$$

We can apply the two-sided Poisson tail bound (Lemma 8) to these three events. We consider only case 1) and case 2) here, because case 3) is eliminated in definition of event E. For the first event, in case 1) we have $\tilde{z}_j(t^*) = 0$. It follows that

$$\begin{aligned}
\mathbb{P}\left(t^*x_j^* - \tilde{z}_j(t^*) > \frac{(T-t^*)(\lambda_j - x_j^*)}{3}\right) &= I\left(t^*x_j^* > \frac{(T-t^*)(\lambda_j - x_j^*)}{3}\right) \\
&= I\left(\frac{(T+2t^*)x_j^*}{3} > \frac{(T-t^*)\lambda_j}{3}\right) \\
&\leq I\left(\frac{3T\lambda_j T^{-1/4}}{3} > \frac{T^{5/6}\lambda_j}{3}\right) \\
&= I\left(T^{3/4}\lambda_j > \frac{T^{5/6}\lambda_j}{3}\right).
\end{aligned}$$

Choosing a constant T_0 such that $T^{5/6}/3 \geq T^{3/4}$ for $T \geq T_0$, we have

$$\mathbb{P}\left(t^*x_j^* - \tilde{z}_j(t^*) > \frac{(T-t^*)(\lambda_j - x_j^*)}{3}\right) \leq 0. \tag{A.42}$$

In case 2), $\tilde{z}_j(t^*)$ is Poisson random variable with parameter $t^*x_j^*$, thus we have

$$\begin{aligned}
&\mathbb{P}\left(t^*x_j^* - \tilde{z}_j(t^*) > \frac{(T-t^*)(\lambda_j - x_j^*)}{3}\right) \\
&\quad \leq 2\exp\left(-\frac{1}{2}\left(\frac{(T-t^*)(\lambda_j - x_j^*)}{3}\right)^2 \frac{1}{(t^*x_j^* + (T-t^*)(\lambda_j - x_j^*)/3)}\right) \\
&\quad \leq 2\exp\left(-\frac{(T-t^*)^2(\lambda_j - x_j^*)^2}{24T\lambda_j}\right) \\
&\quad \leq 2\exp\left(-\frac{(T^{5/6})^2(\lambda_j T^{-1/4})^2}{24T\lambda_j}\right) \\
&\quad \leq 2\exp\left(-\frac{\lambda_{min}}{24}T^{1/6}\right).
\end{aligned} \tag{A.43}$$

For the second term, we have

$$
\begin{aligned}
&\mathbb{P}\left(\Gamma(T) > \frac{(T-t^*)(\lambda_j - x_j^*)}{3}\right) \\
&= \mathbb{P}\left(\alpha \sum_{j\in J_\lambda} |\Lambda_j(T) - \lambda_j T| > \frac{(T-t^*)(\lambda_j - x_j^*)}{3}\right) \\
&\le \sum_{j\in J_\lambda} \mathbb{P}\left(|\Lambda_j(T) - \lambda_j T| > \frac{(T-t^*)(\lambda_j - x_j^*)}{3\alpha|J_\lambda|+1}\right) \\
&\le \sum_{j\in J_\lambda} 2\exp\left(-\frac{1}{2}\left(\frac{(T-t^*)(\lambda_j - x_j^*)}{3\alpha|J_\lambda|+1}\right)^2 \frac{1}{(\lambda_j T + (T-t^*)(\lambda_j - x_j^*)/(3\alpha|J_\lambda|+1))}\right) \\
&\le 2\exp\left(-\frac{1}{2}\frac{((T-t^*)(\lambda_j - x_j^*))^2}{3\alpha|J_\lambda|+1}\frac{1}{(3\alpha|J_\lambda|+2)\lambda_j T}\right) \\
&\le \sum_{j\in J_\lambda} 2\exp\left(-\frac{(T^{5/6}\cdot\lambda_j T^{-1/4})^2}{2(3\alpha|J_\lambda|+1)(3\alpha|J_\lambda|+2)\lambda_j T}\right) \\
&\le 2|J_\lambda|\exp\left(-\frac{\lambda_{min}}{2(3\alpha|J_\lambda|+1)(3\alpha|J_\lambda|+2)}T^{1/6}\right).
\end{aligned}
\tag{A.44}
$$

Moreover, for the last term we have

$$
\begin{aligned}
&\mathbb{P}\left(|\Delta_j(t^*)| > \frac{(T-t^*)(\lambda_j - x_j^*)}{3}\right) \\
&= \mathbb{P}\left(|\Lambda_j(T) - \Lambda_j(t^*) - \lambda_j(T-t^*)| > \frac{(T-t^*)(\lambda_j - x_j^*)}{3}\right) \\
&\le 2\exp\left(-\frac{1}{2}\left(\frac{(T-t^*)(\lambda_j - x_j^*)}{3}\right)^2 \frac{1}{(\lambda_j(T-t^*) + (T-t^*)(\lambda_j - x_j^*)/3)}\right) \\
&\le 2\exp\left(-\frac{T^{5/6}(\lambda_j T^{-1/4})^2}{24\lambda_j}\right) \\
&\le 2\exp\left(-\frac{\lambda_{min}}{24}T^{1/3}\right).
\end{aligned}
\tag{A.45}
$$

From union bound, it follows from (A.42), (A.43), (A.44) and (A.45) that

$$
\mathbb{P}(E_{2,j}^c) \le O\left(\exp\left(-\frac{\lambda_{min}}{24(\alpha|J_\lambda|+1)^2}T^{1/6}\right)\right). \tag{A.46}
$$

Let

$$\kappa = \frac{\lambda_{min}}{24(\alpha|J_\lambda|+1)^2},$$

The results from (A.36), (A.41) and (A.46) lead to

$$\mathbb{P}(E^c) = O\left(n|J_\lambda|\exp(-\kappa T^{1/6})\right),$$

where the big O notation hides an absolute constant. $\square$

Lemma 11. *Probabilities of the following events are bounded by*

$$\mathbb{P}(Q_1) \geq 0.0013 - \frac{0.9496}{\sqrt{T'}}, \quad \mathbb{P}(Q_2) \geq 0.5, \quad \mathbb{P}(Q_3) \geq 9.8531 \times 10^{-10} - \frac{0.9496}{\sqrt{T'}}$$
$$\mathbb{P}(B^c \cap B_1^c \cap Q_1 \cap Q_2) \leq e^{-0.0026\sqrt{T'}},$$

where the events Q_1, Q_2, Q_3, B and B_1 are defined in (A.21)–(A.25) *respectively.*

Proof of Lemma 11. We will apply Berry-Esseen theorem (see formal statement in Lemma 4) to bound $\mathbb{P}(Q_1)$ and $\mathbb{P}(Q_3)$. The probability $\mathbb{P}(Q_1)$ can be written as

$$\begin{aligned}
\mathbb{P}(Q_1) &= \mathbb{P}(T' - 4\sqrt{T'} \leq \Lambda_1(0,T') \leq T' - 3\sqrt{T'}) \\
&= \mathbb{P}\left(-4 \leq \frac{\Lambda_1(0,T') - T'}{\sqrt{T'}} \leq -3\right) \\
&= F_{T'}(-3) - F_{T'}(-4) \\
&= \Phi(-3) - \Phi(-4) - (\Phi(-3) - F_{T'}(-3)) - (F_{T'}(-4) - \Phi(-4)) \\
&\geq \Phi(-3) - \Phi(-4) - |F_{T'}(-3) - \Phi(-3)| - |F_{T'}(-4) - \Phi(-4)|, \qquad \text{(A.47)}
\end{aligned}$$

where $F_{T'}$ is a CDF of $\frac{\Lambda_1(0,T')-T'}{\sqrt{T'}}$. Recall that by the stationary and independent increment properties of Poisson processes, $\Lambda_1(0,T') - T'$ can be thought as the summation of T' i.i.d. $\Lambda_1(1) - 1$ random variables with $\mathbb{E}[\Lambda_1(1) - 1] = 0$, $\mathbb{E}[(\Lambda_1(1) - 1)^2] = 1$ and $\mathbb{E}[|\Lambda_1(1) - 1|^3] = 1$. Hence, the second and the third term of (A.47) can be bounded by the Berry-

Esseen theorem (Lemma 4). That is, for any x, we have

$$|F_{T'}(x) - \Phi(x)| \leq \frac{0.4748}{\sqrt{T'}}. \tag{A.48}$$

Combining (A.48) to (A.47), we can conclude that

$$\mathbb{P}(Q_1) \geq 0.0013 - \frac{0.9496}{\sqrt{T'}}. \tag{A.49}$$

We will apply the same argument to bound the probablity $\mathbb{P}(Q_3)$. That is, we have

$$\begin{aligned}
\mathbb{P}(Q_3) &= \mathbb{P}(T' + 6\sqrt{T'} \leq \Lambda_1(T'', T) \leq T' + 7\sqrt{T'}) \\
&= \mathbb{P}(6 \leq \frac{\Lambda_1(T'', T) - T'}{\sqrt{T'}} \leq 7) \\
&= F_{T'}(7) - F_{T'}(6) \\
&\geq \Phi(7) - \Phi(6) - 2\frac{0.4748}{\sqrt{T'}} \qquad \text{(A.50)} \\
&= 9.8531 \times 10^{-10} - \frac{0.9496}{\sqrt{T'}}, \qquad \text{(A.51)}
\end{aligned}$$

where (A.50) follows from the Lemma 4. Next, to bound the probability $\mathbb{P}(Q_2)$, we will apply Doob's maximal inequality (Lemma 5). Recall that the event Q_2 is defined as

$$Q_2 = \{(t - T') - 2\sqrt{T'} \leq \Lambda_1(T', t) \leq (t - T') + 2\sqrt{T'},\ \forall t \in (T', T'']\}.$$

Equivalently, this event can be re-written as

$$Q_2 = \{\sup_{t \in (T', T'']} |\Lambda_1(T', t) - (t - T')| \leq 2\sqrt{T'}\}.$$

Hence, we have

$$\mathbb{P}(Q_2) = \mathbb{P}(\sup_{t\in(T',T'']} |\Lambda_1(T',t) - (t - T')| \leq 2\sqrt{T'})$$
$$= 1 - \mathbb{P}(\sup_{t\in(T',T'']} |\Lambda_1(T',t) - (t - T')| > 2\sqrt{T'}). \tag{A.52}$$

To bound the last term of (A.52), we first observe that $\Lambda_1(T',t) - (t - T')$ is a martingale with paths that are right continuous with left limits, so we can apply Doob's maximal inequality (Lemma 5) to (A.53), and then use Cauchy-Schwarz inequality in (A.54). We have

$$\mathbb{P}(\sup_{t\in(T',T'']} |\Lambda_1(T',t) - (t - T')| > 2\sqrt{T'}) \leq \frac{\mathbb{E}[|\Lambda_1(T',T'') - (T'' - T')|]}{2\sqrt{T'}} \tag{A.53}$$
$$\leq \frac{\sqrt{\mathbb{E}[(\Lambda_1(T',T'') - (T'' - T'))^2]}}{2\sqrt{T'}} \tag{A.54}$$
$$= \frac{\sqrt{\mathrm{Var}(\Lambda_1(T',T''))}}{2\sqrt{T'}} = \frac{\sqrt{T'}}{2\sqrt{T'}}$$
$$= 0.5. \tag{A.55}$$

Combining the result in (A.55) to (A.52), we get

$$\mathbb{P}(Q_2) \geq 1 - 0.5 = 0.5. \tag{A.56}$$

Next, we will bound the probability $\mathbb{P}(B^c \cap B_1^c \cap Q_1 \cap Q_2)$. If the event $B^c \cap B_1^c \cap Q_1 \cap Q_2$ happens, we can observe that $z_1(0,T') \leq \Lambda_1(0,T') \leq T' - 3\sqrt{T'}$ from Q_1 and $z_2(0,T') \leq z_2(0,T) < \frac{1}{6}\sqrt{T'}$ from B^c, and hence the remaining capacity at time T' will be

$$C(T') = C - z_1(0,T') - z_2(0,T') \geq T - (T' - 3\sqrt{T'}) - \frac{1}{6}\sqrt{T'} = T - T' + \frac{17}{6}\sqrt{T'}.$$

Similarly, for any time $t \in (T', T'']$, the event $B^c \cap B_1^c \cap Q_1 \cap Q_2$ implies that

$$z_1(0,t) \leq \Lambda_1(0,T') + \Lambda_1(T',t) \leq T' - 3\sqrt{T'} + (t - T) + 2\sqrt{T'}$$

from Q_1 and Q_2 and $z_2(0,t) \leq z_2(0,T) < \frac{1}{6}\sqrt{T'}$ from B^c; therefore, the remaining capacity at any time $t \in (T', T'']$ will be

$$\begin{aligned} C(t) = C - z_1(0,t) - z_2(0,t) &\geq T - (T' - 3\sqrt{T'}) - ((t - T') + 2\sqrt{T'}) - \frac{1}{6}\sqrt{T'} \\ &= T - t + \frac{5}{6}\sqrt{T'}. \end{aligned}$$

So the average capacity per period at time T' is

$$b(T') \geq \frac{T - T' + \frac{17}{6}\sqrt{T'}}{T - T'} = 1 + \frac{17\sqrt{T'}/6}{2T'} \geq 1 + \frac{17}{12\sqrt{T'}}, \tag{A.57}$$

and similarly the average capacity per period at any time $t \in (T', T'']$ is given by

$$b(t) \geq \frac{T - t + \frac{5}{6}\sqrt{T'}}{T - t} = 1 + \frac{5\sqrt{T'}/6}{T - t} \geq 1 + \frac{5}{12\sqrt{T'}}. \tag{A.58}$$

Recall that for the problem instance we consider, the admission probability of class 1 customer, which is obtained from the LP described in Algorithm 2, is given by $x_1(t) = \min(b(t)/\lambda_1, 1) = \min(b(t), 1)$. Thus, (A.57) and (A.58) implies that the decision maker must accept all arrivals of class 1 customer in phase 2. Hence, it follows from definition of the event Q_2 in (A.22) that, for any time $t \in (T', T'']$, we have

$$(t - T') - 2\sqrt{T'} \leq z_1(t, T') \leq (t - T') + 2\sqrt{T'}. \tag{A.59}$$

Next, B_1^c implies that we can upper bound the remaining capacity at time T' as

$$C(T') = C - z_1(0,T') - z_2(0,T') \leq T - z_1(0,T') \leq T - (T' - 10\sqrt{T'}) = T - T' + 10\sqrt{T'},$$

which results in

$$b(T') \leq \frac{T - T' + 10\sqrt{T'}}{T - T'} = 1 + \frac{10\sqrt{T'}}{T - T'} \leq 1 + \frac{5}{\sqrt{T'}}. \tag{A.60}$$

Similarly, we can upper bound the remaining capacity at time $t \in (T', T'']$ using the results in (A.59) which yields

$$\begin{aligned} C(t) &= C - z_1(0,t) - z_2(0,t) \leq T - z_1(0,T') - z_1(T',t) \\ &\leq T - (T' - 10\sqrt{T'}) - ((t - T') - 2\sqrt{T'}) = T - t + 12\sqrt{T'}. \end{aligned}$$

It follows that the upper bound of the average capacity per period at time $t \in (T', T'']$ is given by

$$b(t) \leq \frac{T - t + 12\sqrt{T'}}{T - t} = 1 + \frac{12\sqrt{T'}}{T - t} \leq 1 + \frac{12}{\sqrt{T'}}. \tag{A.61}$$

Combining the results from (A.57), (A.58), (A.60) and (A.61), we obtain the bound of the average capacity per period at $t \in [T', T'']$, that is,

$$1 + \frac{5}{12\sqrt{T'}} \leq b(t) \leq 1 + \frac{12}{\sqrt{T'}}, \tag{A.62}$$

which also implies that the solution to the LP at time $t \in [T', T'']$ satisfies $\frac{5}{12\sqrt{T'}} \leq x_2(t) \leq \frac{12}{\sqrt{T'}}$. Therefore, the probability of the event $B^c \cap B_1^c \cap Q_1 \cap Q_2$ can be written as

$$\mathbb{P}(B^c \cap B_1^c \cap Q_1 \cap Q_2) \leq \mathbb{P}\left(\frac{5}{12\sqrt{T'}} \leq x_2(t) \leq \frac{12}{\sqrt{T'}}, \forall t \in [T', T''],\ z_2(T', T'') < \frac{1}{6}\sqrt{T'}\right). \tag{A.63}$$

We will use Freedman's Inequality (Lemma 6) to bound (A.63). For $i \in [T']$, we let $\xi_0 = 0$ and $\xi_i = x_2(T' + i - 1) - z_2(T' + i - 1, T' + i)$. We observe that ξ_i is $\mathscr{F}_{T'+i-1}$-measurable and $\mathbb{E}[\xi_i | \mathscr{F}_{T'+i-1}] = x_2(T' + i - 1) - x_2(T' + i - 1) = 0$. Thus, ξ_i is a martingale difference.

Set

$$S_{T'} = \sum_{i=1}^{T'} \xi_i = \sum_{i=1}^{T'} x_2(T'+i-1) - z_2(T',T'').$$

Then, conditions in (A.63) imply that

$$S_{T'} > \frac{5}{12\sqrt{T'}}T' - \frac{1}{6}\sqrt{T'} = \frac{1}{4}\sqrt{T'}.$$

Moreover, let $\langle S \rangle_k = \sum_{i=1}^{k} \mathbb{E}[\xi_i^2 | \mathscr{F}_{i-1}]$ for $k \geq 1$. We have

$$\begin{aligned}
\langle S \rangle_{T'} &= \sum_{i=1}^{T'} \mathbb{E}[\xi_i^2 | \mathscr{F}_{T'+i-1}] = \sum_{i=1}^{T'} \mathbb{E}[(x_2(T'+i-1) - z_2(T'+i-1, T'+i))^2 | \mathscr{F}_{T'+i-1}] \\
&= \sum_{i=1}^{T'} \mathrm{Var}(z_2(T'+i-1, T'+i) | \mathscr{F}_{T'+i-1}) = T' x_2(T'+i-1).
\end{aligned}$$

Then, conditions in (A.63) imply that

$$\langle S \rangle_{T'} \leq T' \frac{12}{\sqrt{T'}} = 12\sqrt{T'}.$$

Therefore, we have

$$\begin{aligned}
&\mathbb{P}\left(\frac{5}{12\sqrt{T'}} \leq x_2(t) \leq \frac{12}{\sqrt{T'}}, \forall t \in [T', T''],\, z_2(T', T'') < \frac{1}{6}\sqrt{T'} \right) \\
&\qquad\qquad \leq \mathbb{P}\left(S_{T'} > \frac{1}{4}\sqrt{T'}, \langle S \rangle_{T'} \leq 12\sqrt{T'} \right). \qquad \text{(A.64)}
\end{aligned}$$

Since, for all $i \in [T']$, we have $\xi_i \leq x_2(T+i-1) \leq \lambda_2 = 1$, we can apply Freedman's inequality (Lemma 6) and get

$$\begin{aligned}
\mathbb{P}\left(S_{T'} > \frac{1}{4}\sqrt{T'}, \langle S \rangle_{T'} \leq 12\sqrt{T'} \right) &\leq \exp\left(-\frac{(\sqrt{T'}/4)^2}{2(12\sqrt{T'} + \sqrt{T'}/4)} \right) \\
&= \exp\left(-\frac{1}{392}\sqrt{T'} \right) = e^{-0.0026\sqrt{T'}}.
\end{aligned}$$

□

Lemma 12 (Bound on b_l)**.** *We have* $\mathbb{E}[(b_l - b_l(t))^+] \le K_l \sqrt{\sum_{i=0}^{t-1} \frac{1}{(T-i-1)^2}}$, *where* $K_l = \sqrt{\sum_{j=1}^{n} a_{lj}^2 \lambda_j^2}$.

Proof of Lemma 12. Recall from (A.32) that we have

$$(b_l - b_l(t))^+ \le \left[\sum_{i=0}^{t-1} \left(\frac{\sum_{j=1}^{n} a_{lj}(\tilde{z}_j(i+1) - \tilde{z}_j(i))}{T-i-1} - \frac{\sum_{j=1}^{n} a_{lj} x_j(i)}{T-i-1}\right)\right]^+.$$

Taking expectations on both sides yields

$$\begin{aligned}
\mathbb{E}[(b_l - b_l(t))^+] &\le \mathbb{E}\left[\left(\sum_{i=0}^{t-1} \frac{\sum_{j=1}^{n} a_{lj}(\tilde{z}_j(i+1) - \tilde{z}_j(i)) - \sum_{j=1}^{n} a_{lj} x_j(i)}{T-i-1}\right)^+\right] \\
&\le \mathbb{E}\left[\left|\sum_{i=0}^{t-1} \frac{\sum_{j=1}^{n} a_{lj}(\tilde{z}_j(i+1) - \tilde{z}_j(i)) - \sum_{j=1}^{n} a_{lj} x_j(i)}{T-i-1}\right|\right] &\text{(A.65)} \\
&\le \sqrt{\mathbb{E}\left[\left(\sum_{i=0}^{t-1} \frac{\sum_{j=1}^{n} a_{lj}(\tilde{z}_j(i+1) - \tilde{z}_j(i)) - \sum_{j=1}^{n} a_{lj} x_j(i)}{T-i-1}\right)^2\right]}. &\text{(A.66)}
\end{aligned}$$

The last line applies the Cauchy-Schwarz inequality.

Let $\{\mathscr{F}_t, 0 \le t \le T\}$ be the filtration generated by $\{\Lambda_j(s), 0 \le s \le t, j \in [n]\}$. By the law of total expectation, (A.66) becomes

$$\mathbb{E}[(b_l - b_l(t))^+] \le \sqrt{\mathbb{E}\left[\mathbb{E}\left[\left(\sum_{i=0}^{t-1} \frac{\sum_{j=1}^{n} a_{lj}(\tilde{z}_j(i+1) - \tilde{z}_j(i)) - \sum_{j=1}^{n} a_{lj} x_j(i)}{T-i-1}\right)^2 \middle| \mathscr{F}_i\right]\right]}. \tag{A.67}$$

Since the conditional independence of the arrivals of the customers in different period implies that the arrivals of the admitted customers in different period are also conditionally independent, and each of the summands has mean zero, and the cross-terms vanish. Thus, (A.67) equals to

$$\mathbb{E}[(b_l - b_l(t))^+] \leq \sqrt{\mathbb{E}\left[\mathbb{E}\left[\sum_{i=0}^{t-1}\left(\frac{\sum_{j=1}^n a_{lj}(\tilde{z}_j(i+1)-\tilde{z}_j(i)) - \sum_{j=1}^n a_{lj}x_j(i)}{T-i-1}\right)^2 \middle| \mathscr{F}_i\right]\right]}.$$

From the description of the FR, we know that $\tilde{z}_j(t+1)-\tilde{z}_j(t)$ conditioned on $\mathscr{F}_t$ is distributed as Poisson distribution with parameter $x_j(t)$ (from the Poisson thinning property). We use the definition of variance in Equation (A.68), and we use the observation that $x(t)$ is bounded above by λ_j because $x(t)$ is the solution to the LP in (A.69). It immediately follows that $\mathbb{E}[\tilde{z}_j(t+1)-\tilde{z}_j(t)|\mathscr{F}_t] = x_j(t)$ and $\mathrm{Var}(\tilde{z}_j(t+1)-\tilde{z}_j(t)|\mathscr{F}_t) = x_j(t)$. Therefore, we can write

$$\begin{aligned}
&\mathbb{E}[(b_l - b_l(t))^+] \\
&\leq \sqrt{\mathbb{E}\left[\sum_{i=0}^{t-1}\frac{1}{(T-i-1)^2}\mathbb{E}\left[\left(\sum_{j=1}^n a_{lj}(\tilde{z}_j(i+1)-\tilde{z}_j(i)) - \sum_{j=1}^n a_{lj}x_j(i)\right)^2 \middle| \mathscr{F}_i\right]\right]} \\
&= \sqrt{\mathbb{E}\left[\sum_{i=0}^{t-1}\frac{1}{(T-i-1)^2}\mathrm{Var}\left(\sum_{j=1}^n a_{lj}(\tilde{z}_j(i+1)-\tilde{z}_j(i)) \middle| \mathscr{F}_i\right)\right]} \qquad \text{(A.68)} \\
&= \sqrt{\sum_{i=0}^{t-1}\frac{1}{(T-i-1)^2}\sum_{j=1}^n a_{lj}^2\,\mathbb{E}\left[\mathrm{Var}\left(\tilde{z}_j(i+1)-\tilde{z}_j(i)|\mathscr{F}_i\right)\right]} \\
&= \sqrt{\sum_{i=0}^{t-1}\frac{1}{(T-i-1)^2}\sum_{j=1}^n a_{lj}^2\,\mathbb{E}\left[x_j(t)^2\right]} \\
&\leq \sqrt{\sum_{i=0}^{t-1}\frac{1}{(T-i-1)^2}\sum_{j=1}^n a_{lj}^2\lambda_j^2} \qquad \text{(A.69)} \\
&= K_l\sqrt{\sum_{i=0}^{t-1}\frac{1}{(T-i-1)^2}},
\end{aligned}$$

where $K_l = \sqrt{\sum_{j=1}^n a_{lj}^2\lambda_j^2}$. □

A.5 Numerical Performance: Comparison with Algorithm proposed by Vera and Banerjee (2021)

In this section, we compare the numerical performance of SPA, FR, IRT, IR and FRT with Fluid Bayes Selector (FBS), which is recently proposed by Vera and Banerjee (2021). The FBS re-solves the DLP when there is an arrival and accepts the arrival if the acceptance probability exceeds 0.5. Instead of re-solving the DLP every time when there is an arrival, we slightly modify re-solving schedule of FBS. Specifically, we divide the horizon into T periods and let FBS re-solves the DLP at the beginning of each period. The complete definition of FBS is given in Algorithm 11.

Algorithm 11 Fluid Bayes Selector: FBS

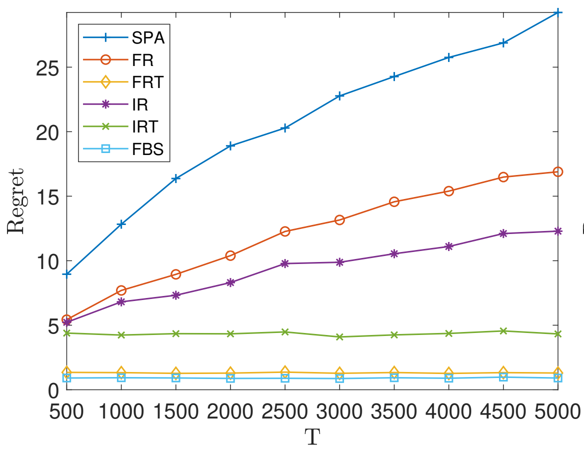
initialize: set $C(0) = C$ and $b(0) = C/T$
for $t = 0, 1, \ldots, T-1$ **do**
  set $x(t) \leftarrow \arg\max_x \left\{ \sum_{j=1}^n r_j x_j \;\middle|\; \sum_{j=1}^n A_j x_j \leq b(t), \text{ and } 0 \leq x_j \leq \lambda_j, \forall j \in [n] \right\}$
  set $C' \leftarrow C(t)$
  for all customers arriving in $[t, t+1)$ **do**
    if the customer belongs to class j and $A_j \leq C'$ ($\forall j \in [n]$) **then**
      if $x_j(t) < \lambda_j/2$ **then**
        reject the customer
      else
        accept the customer
      if the customer is accepted, update $C' \leftarrow C' - A_j$
    else
      reject the customer
  set $C(t+1) \leftarrow C'$ and $b(t+1) \leftarrow \frac{C(t+1)}{T-t-1}$

We consider the NRM with a single resource and two classes of customers (the same setting as Sec. 6.1). Recall that the arrival process of each class of customer follows an independent Poisson process with rate 1. Customers from both classes, if accepted, consume one unit of resource, but pay different prices, r_1 and r_2.

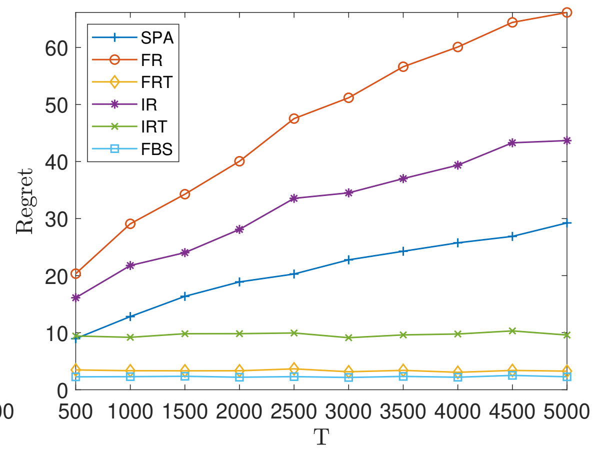
Figure A.1 plots the regret under SPA, FR, FRT, IR, IRT and FBS over 1000 sample paths. The first column shows the case when $r_1 = 2$ and $r_2 = 1$, while the second column shows the case when $r_1 = 5$ and $r_2 = 1$. The first, the second and the third rows illustrate the case when $b = 1$, $b = 1.1$, and $b = 1.5$ respectively. Figure A.2 plots the regret under FR, IRT and the FBS when $r_1 = 2, r_2 = 1$ and $b = 0.5, 0.6, \ldots, 2$ for $T = 50000$ over 1000 sample paths.

It can be observed from Figure A.1 and Figure A.2 that the regret of FBS remains constant regardless of the horizon length or the average capacity per unit time, which is consistent with the theoretical results in Vera and Banerjee (2021). In fact, the regret of FBS is the smallest among the algorithms we consider.

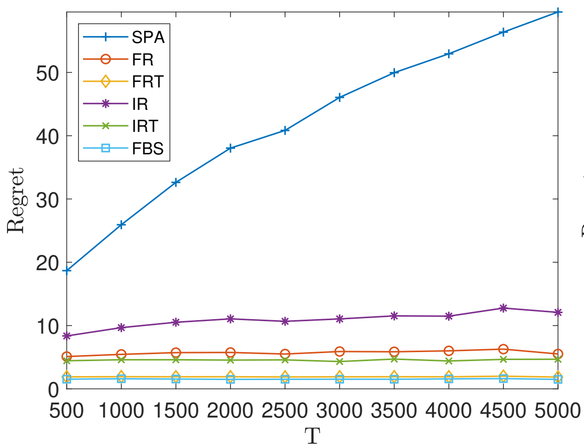
Because IRT and FBS are proven to have $O(1)$ revenue loss, we further investigate the actual decisions made by these two algorithms. We compare the solutions obtained from IRT and FBS when $r_1 = 2, r_2 = 1, b = 1$ and $T = 250$ over 1000 sample paths. Figure A.3 shows the probability that different decisions made by IRT and FBS (accept vs reject or reject vs accept) in each time period when $T = 250$. Vertical dashed lines indicate re-solving time under IRT. It can be observed that before the first re-solving time of IRT, both algorithms perform almost exactly the same, that is, accept all class 1 customers and reject all class 2 customers. Then, IRT and FBS start making slightly different decisions after the first re-solving time of IRT. Also, we can observe a jump in probability of different decisions made by these two algorithms after each re-solving time of IRT.



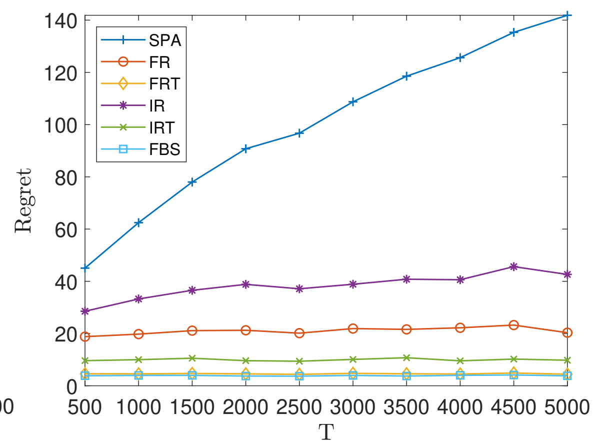
(a) $b = 1, r_1 = 2$ and $r_2 = 1$



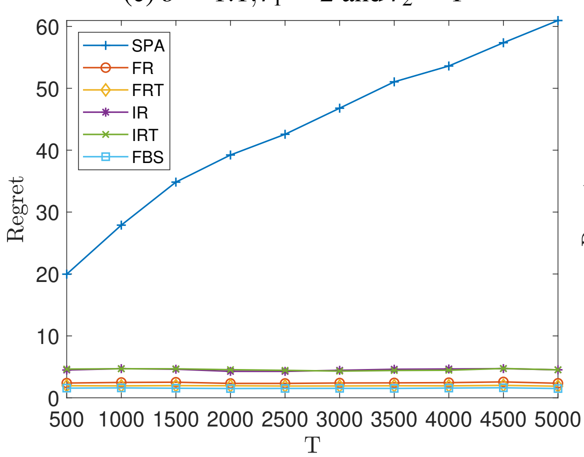
(b) $b = 1, r_1 = 5$ and $r_2 = 1$



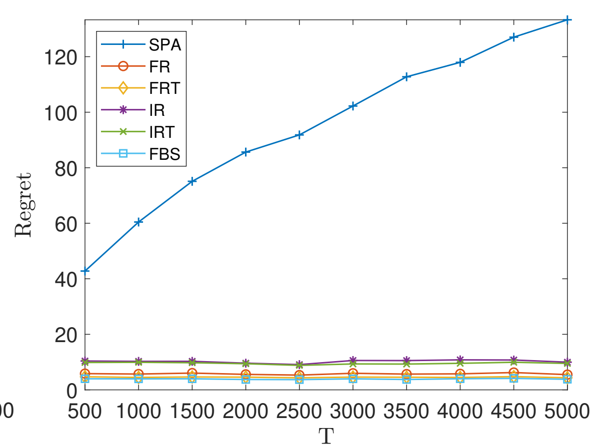
(c) $b = 1.1, r_1 = 2$ and $r_2 = 1$



(d) $b = 1.1, r_1 = 5$ and $r_2 = 1$



(e) $b = 1.5, r_1 = 2$ and $r_2 = 1$



(f) $b = 1.5, r_1 = 5$ and $r_2 = 1$

Figure A.1: Regret under the SPA, the FR, the FRT, the IR, the IRT, and the FBS policies for $T = 500, 1000, \ldots, 5000$.

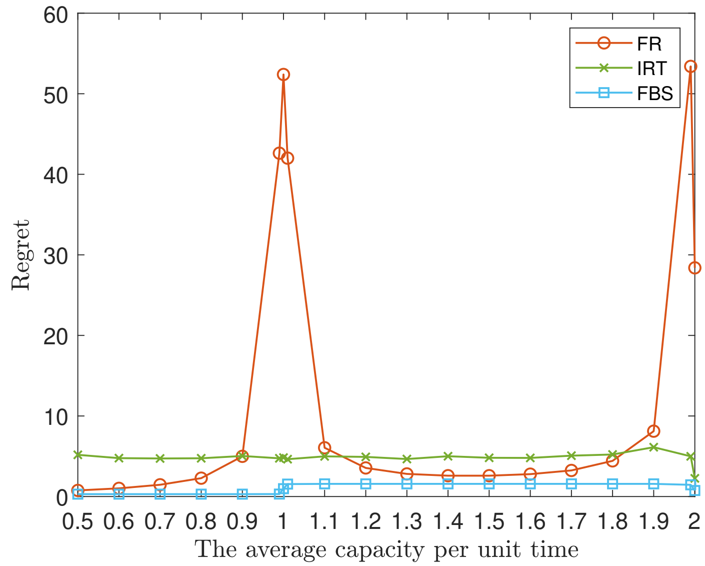
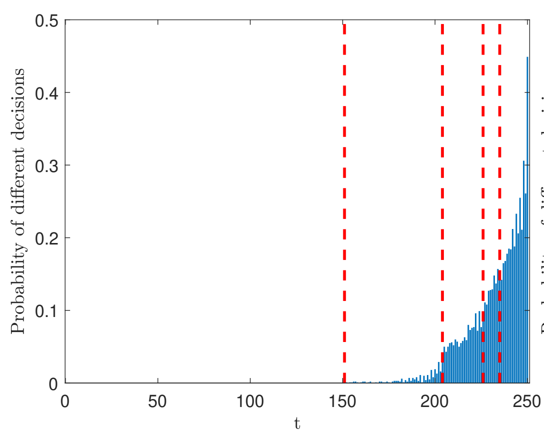
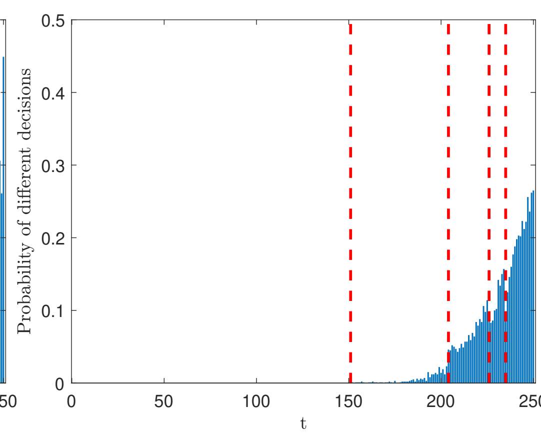


Figure A.2: Regret under the FR policy, the IRT policy and FBS policy for $r_1 = 2, r_2 = 1$ and $T = 50000$.



(a) class 1 customer



(b) class 2 customer

Figure A.3: Probability of different decisions made under IRT and FBS in each period for $T = 250$.

APPENDIX B

INTEGRATED PRICING AND ROUTING IN A NETWORK

B.1 Arc-based Formulation

Let $v_{j,e}$ be the demand of OD pair j on arc e for all $j \in \mathscr{J}$ and $e \in \mathscr{E}$. Let S_j and D_j denote the origin and destination associated with OD pair j for all $j \in \mathscr{J}$. Furthermore, let $\mathscr{E}^-(i)$ and $\mathscr{E}^+(i)$ denote the set of incoming and outgoing arcs of node i, respectively. The integrated pricing and routing problem can also be formulated as

$$\max_{v,x,z} \quad \sum_{j \in \mathscr{J}} r_j(z_j) - \sum_{e \in \mathscr{E}} c_e v_e \tag{B.1a}$$

$$\text{s.t.} \quad \sum_{e \in \mathscr{E}^+(i)} v_{j,e} - \sum_{e \in \mathscr{E}^-(i)} v_{j,e} = \begin{cases} z_j & \text{for } i = S_j \\ 0 & \text{for } i \in \mathscr{N}, i \neq S_j, D_j, \\ -z_j & \text{for } i = D_j \end{cases} \forall j \in \mathscr{J}, \tag{B.1b}$$

$$v_e = \sum_{j \in \mathscr{J}} v_{j,e}, \qquad \forall e \in \mathscr{E}, \tag{B.1c}$$

$$v_e \leq k_e, \qquad \forall e \in \mathscr{E}, \tag{B.1d}$$

$$v_e \geq 0, \qquad \forall e \in \mathscr{E}. \tag{B.1e}$$

Constraints (B.1b) ensure demand is delivered from origin S_j to destination D_j for all OD pairs $j \in \mathscr{J}$. Constraints (B.1c) determine the total flow on each arc in the network. Constraints (B.1d) enforce the capacity restrictions for the arcs. Constraints (B.1e) ensure that the demand allocated to an arc is non-negative. The arc-based formulation has $|\mathscr{J}| + |\mathscr{E}| + |\mathscr{J}||\mathscr{E}|$ decision variables and $|\mathscr{N}||\mathscr{J}| + 2|\mathscr{E}|$ constraints.

B.2 Proofs of Theorems

B.2.1 Proof of Theorem 3

Proof of Theorem 3. Define $\tilde{L}(z,y) = \max_{\substack{z\geq 0 \\ z_j = \sum_{p\in\mathscr{P}_j} x_p}} L(x,y)$. It follows that

$$\begin{aligned}
\max_{z\geq 0} \sum_{t=1}^{T} \tilde{L}(z,y_t) &= \max_{z\geq 0} \sum_{t=1}^{T} \max_{\substack{z\geq 0 \\ z_j = \sum_{p\in\mathscr{P}_j} x_p}} L(x,y_t) \\
&= \max_{z\geq 0} \max_{\substack{z\geq 0 \\ z_j = \sum_{p\in\mathscr{P}_j} x_p}} \sum_{t=1}^{T} L(x,y_t) \\
&= \max_{x\geq 0} \sum_{t=1}^{T} L(x,y_t).
\end{aligned}$$

Because we apply Online Supergradient Ascent on a sequence of $\tilde{L}(z,y_t)$ which is strongly concave in z, from Lemma 14, we have

$$\begin{aligned}
O(\log T) &\geq \max_{z\geq 0} \sum_{t=1}^{T} \tilde{L}(z,y_t) - \sum_{t=1}^{T} \tilde{L}(z_t,y_t) \\
&= \max_{x\geq 0} \sum_{t=1}^{T} L(x,y_t) - \sum_{t=1}^{T} \tilde{L}(z_t,y_t).
\end{aligned} \tag{B.2}$$

We have shown previously that x and y which solve MCMCF in (3.11) are identical to x and y which solve $\min_{y\geq 0} \max_{\substack{x\geq 0 \\ \sum_{p\in\mathscr{P}_j} x_p = z_j}} L(x,y)$. Therefore, from definition of x_t and y_t in Algorithm 8, we have

$$\begin{aligned}
L(x_t,y_t) = \min_{y\geq 0} \max_{\substack{x\geq 0 \\ \sum_{p\in\mathscr{P}_j} x_p = z_{j,t}}} L(x,y) &= \max_{\substack{x\geq 0 \\ \sum_{p\in\mathscr{P}_j} x_p = z_{j,t}}} \min_{y\geq 0} L(x,y) \\
= \min_{y\geq 0} L(x_t,y) &= \max_{\substack{x\geq 0 \\ \sum_{p\in\mathscr{P}_j} x_p = z_{j,t}}} L(x,y_t).
\end{aligned} \tag{B.3}$$

By definition of $\tilde{L}$, we know that

$$\tilde{L}(z_t, y_t) = \max_{\substack{x \geq 0 \\ \sum_{p \in \mathscr{P}_j} x_p = z_{j,t}}} L(x, y_t) = L(x_t, y_t). \tag{B.4}$$

Hence, from (B.2), it follows that

$$\begin{aligned} O(\log T) &\geq \max_{x \geq 0} \sum_{t=1}^{T} L(x, y_t) - \sum_{t=1}^{T} L(x_t, y_t) \\ &\geq \sum_{t=1}^{T} L(x^*, y_t) - \sum_{t=1}^{T} L(x_t, y_t) \qquad \text{(B.5)} \\ &\geq T L(x^*, \sum_{t=1}^{T} y_t / T) - \sum_{t=1}^{T} L(x_t, y_t) \qquad \text{(B.6)} \\ &\geq T L(x^*, y^*) - \sum_{t=1}^{T} L(x_t, y_t), \qquad \text{(B.7)} \end{aligned}$$

where (B.5) follows the feasibility of x^*, (B.6) follows the convexity in y, and (B.7) follows the definition of y^*. Therefore, we have

$$L(x^*, y^*) - \frac{1}{T} \sum_{t=1}^{T} L(x_t, y_t) \leq O\left(\frac{\log T}{T}\right). \tag{B.8}$$

From (B.3), we know that $y_t = \arg\min_{y \geq 0} L(x_t, y)$. It follow from Lemma 15 that

$$\begin{aligned} 0 &\geq \sum_{t=1}^{T} L(x_t, y_t) - \min_{y \geq 0} \sum_{t=1}^{T} L(x_t, y) \\ &\geq \sum_{t=1}^{T} L(x_t, y_t) - L(x_t, y^*) \qquad \text{(B.9)} \\ &\geq \sum_{t=1}^{T} L(x_t, y_t) - T L(\sum_{t=1}^{T} x_t / T, y^*) \qquad \text{(B.10)} \\ &\geq \sum_{t=1}^{T} L(x_t, y_t) - T L(x^*, y^*), \qquad \text{(B.11)} \end{aligned}$$

where (B.9) follows the feasibility of y^*, (B.10) follows the concavity of L in x, and (B.11) follows the definition of x^*. Therefore, we have

$$\frac{1}{T}\sum_{t=1}^{T} L(x_t, y_t) - L(x^*, y^*) \leq 0. \tag{B.12}$$

Combining (B.8) and (B.12), we can conclude that

$$\left|\frac{1}{T}\sum_{t=1}^{T} L(x_t, y_t) - L(x^*, y^*)\right| \leq O\left(\frac{\log T}{T}\right).$$

$\square$

B.2.2 Proof of Theorem 4

Proof of Theorem 4. Let $x_p = \rho_j(p) z_j$ for all $p \in \mathscr{P}_j, j \in \mathscr{J}$, $\sum_{p\in\mathscr{P}_j} \rho_j(p) = 1$ for all $j \in \mathscr{J}$ and $\rho_j(p) \geq 0$ for all $p \in \mathscr{P}_j, j \in \mathscr{J}$. We define $H(z, y^*) = \sum_{j\in\mathscr{J}} r_j(z_j) - \sum_{e\in\mathscr{E}} c_e \sum_{j\in J}\sum_{p\in\mathscr{P}_j} \delta_e(p)\rho_j(p) z_j + \sum_{e\in\mathscr{E}} y_e^*(k_e - \sum_{j\in J}\sum_{p\in\mathscr{P}_j} \delta_e(p)\rho_j(p) z_j)$. That is, we have $H(z, y^*) = L(x, y^*)$. It is easy to verify that H is μ-strongly concave in z, so we have

$$H(z, y^*) \leq H(z^*, y^*) + \nabla_z H(z^*, y^*)^\top (z - z^*) - \frac{\mu}{2}\|z - z^*\|^2.$$

Because z^* is a maximizer, it follows that $\nabla_z H(z^*, y^*)^\top (z - z^*) \leq 0$, and hence

$$H(z, y^*) \leq H(z^*, y^*) - \frac{\mu}{2}\|z - z^*\|^2.$$

Because $L(x, y^*) = H(z, y^*)$, we have

$$T L(x^*, y^*) - \sum_{t=1}^{T} L(x_t, y^*) \geq \frac{\mu}{2}\sum_{t=1}^{T} \|z_t - z^*\|^2.$$

It follows that

$$
\begin{aligned}
TL(x^*,y^*) - \sum_{t=1}^{T} L(x_t,y^*) &= TL(x^*,y^*) - \sum_{t=1}^{T} L(x_t,y_t) + \sum_{t=1}^{T} L(x_t,y_t) - \sum_{t=1}^{T} L(x_t,y^*) \\
&\leq O(\log T),
\end{aligned}
$$

where the inequality follows from (B.7) and (B.11). Therefore,

$$
\begin{aligned}
O(\log T) &\geq \sum_{t=1}^{T} \|z_t - z^*\|^2 \\
&= T\left[\frac{\sum_{t=1}^{T} \|z_t - z^*\|^2}{T}\right] \\
&\geq T\left[\left\|\sum_{t=1}^{T} z_t/T - z^*\right\|^2\right],
\end{aligned}
$$

where the last equation follows from Jensen's inequality because $\|\cdot\|^2$ is convex. Therefore, we have

$$
\|\bar{z} - z^*\|^2 \leq O\left(\frac{\log T}{T}\right).
$$

□

B.3 Lemmas

Lemma 13. *Let $x^* = \arg\max_{x \in X} f(x)$. When f is β-smooth, under Algorithm 5, we have*

$$
f(x^*) - f(x_T) \leq O\left(\frac{1}{T}\right).
$$

Proof of Lemma 13. Since f is β-smooth, we have

$$\begin{aligned}
f(x_{s+1}) &\geq f(x_s) + \nabla f(x_s)^\top (x_{s+1} - x_s) - \frac{\beta}{2}\|x_{s+1} - x_s\|^2 \\
&= f(x_s) + \nabla f(x_s)^\top (x_s + \gamma_s(w_s - x_s) - x_s) - \frac{\beta}{2}\|x_s + \gamma_s(w_s - x_s) - x_s\|^2 \\
&\geq f(x_s) + \gamma_s \nabla f(x_s)^\top (w_s - x_s) - \frac{\beta\gamma_s^2 R^2}{2} \\
&\geq f(x_s) + \gamma_s \nabla f(x_s)^\top (x^* - x_s) - \frac{\beta\gamma_s^2 R^2}{2} \qquad \text{(B.13)} \\
&\geq f(x_s) + \gamma_s (f(x^*) - f(x_s)) - \frac{\beta\gamma_s^2 R^2}{2}, \qquad \text{(B.14)}
\end{aligned}$$

where (B.13) follows the definition of z_s and (B.14) holds because f is concave.

$$f(x^*) - f(x_{s+1}) \leq (1 - \gamma_s)(f(x^*) - f(x_s)) + \frac{\beta\gamma_s^2 R^2}{2}. \qquad \text{(B.15)}$$

Let $\Delta_s = f(x^*) - f(x_s)$. (B.15) can be re-written as

$$\Delta_{s+1} \leq (1 - \gamma_s)\Delta_s + \gamma_s^2 \frac{\beta R^2}{2}.$$

When $\Delta_s = \frac{2}{s+2}$. We can show by induction that $\Delta_s \leq \frac{2\beta R^2}{s+2}$ for $s = 0, 1, \ldots$. For base case ($s = 0$), we have $\gamma_0 = 1$, and hence, $\Delta_1 \leq \frac{\beta R^2}{2} \leq \beta R^2$. Assume that $\Delta_s \leq \frac{2\beta R^2}{s+2}$ holds for $s \geq 0$. We have

$$\begin{aligned}
\Delta_{s+1} &\leq \left(1 - \frac{2}{s+2}\right)\frac{2\beta R^2}{s+2} + \left(\frac{2}{s+2}\right)^2 \frac{\beta R^2}{2} \\
&= \frac{2s\beta R^2}{(s+2)^2} + \frac{2\beta R^2}{(s+2)^2} \\
&= \frac{2\beta R^2}{(s+2)^2}(s+1) \\
&\leq \frac{2\beta R^2}{(s+1)(s+3)}(s+1) \qquad \text{(B.16)} \\
&= \frac{2\beta R^2}{s+3},
\end{aligned}$$

where the inequality (B.16) holds because $(s+2)^2 \geq (s+1)(s+3)$. $\square$

Lemma 14 (Theorem 3.3 in Hazan (2016))**.** *For any sequences of μ-strongly convex functions ℓ_t, Online gradient descent (Algorithm 7) with step sizes $\eta_t = \frac{1}{\mu t}$, it follows that*

$$\sum_{t=1}^{T} \ell_t(x_t) - \min_{x \in X} \sum_{t=1}^{T} \ell_t(x) \leq \frac{G^2}{2\alpha}(1 + \log T),$$

where $\|\nabla f(x)\| \leq G$ for all $x \in X$.

Algorithm 12 Best Response

Input: T
for $t = 0, 1, 2, \ldots, T$ **do**
  Set $x_t \leftarrow \arg\min_{x \in X} \ell_t(x)$

Lemma 15. *For any sequences of convex functions ℓ_t, Best Response (Algorithm 12) achieves*

$$\sum_{t=1}^{T} \ell_t(x_t) - \min_{x \in X} \sum_{t=1}^{T} \ell_t(x) \leq 0.$$

Proof of Lemma 15. From the description of Best Response, for $t = 1, \ldots, T$, we have $\ell_t(x_t) \leq \ell_t(x)$ for all $x \in X$. It follows that

$$\begin{aligned} \sum_{t=1}^{T} \ell_t(x_t) &\leq \sum_{t=1}^{T} \ell_t(x), \quad \forall x \in X \\ &\leq \min_{x \in X} \sum_{t=1}^{T} \ell_t(x). \end{aligned}$$

$\square$

APPENDIX C

DYNAMIC PRICING AND MATCHING IN TWO-SIDED QUEUES

C.1 MDP Analysis

C.1.1 Proof of Proposition 4

Proof of Proposition 4. Since the delay-sensitive model and the holding cost model have the same states, actions, and transition rates, a control policy π induces the same stationary distribution under the two models. Let $(\boldsymbol{\lambda}^{\pi}(\mathbf{q}), \boldsymbol{\mu}^{\pi}(\mathbf{q}))$ be the arrival rates under this policy when the system state is $\mathbf{q}$, and let $\mathbb{E}^{\pi}[\cdot]$ be the expectation operator under the stationary distribution induced by this policy. The long-run average profit of the delay sensitive model is given by

$$
\begin{aligned}
&\mathbb{E}^{\pi}\left[\sum_{j=1}^{m} \lambda_j^{\pi}(\mathbf{q})\left(F_j(\lambda_j^{\pi}(\mathbf{q})) - s_j^{(c)}\,\mathbb{E}^{\pi}[w_j^{(c)}]\right) - \sum_{i=1}^{n} \mu_i^{\pi}(\mathbf{q})\left(G_i(\mu_i^{\pi}(\mathbf{q})) + s_i^{(s)}\,\mathbb{E}^{\pi}[w_i^{(s)}]\right)\right] \\
&= \mathbb{E}^{\pi}\left[\sum_{j=1}^{m} \lambda_j^{\pi}(\mathbf{q}) F_j\left(\lambda_j^{\pi}(\mathbf{q})\right) - \sum_{i=1}^{n} \mu_i^{\pi}(\mathbf{q}) G_i\left(\mu_i^{\pi}(\mathbf{q})\right)\right] - \sum_{j=1}^{m} s_j^{(c)}\,\mathbb{E}[\lambda_j^{\pi}(\mathbf{q})]\,\mathbb{E}^{\pi}[w_j^{(c)}(\mathbf{q})] \\
&\quad - \sum_{i=1}^{n} s_i^{(s)}\,\mathbb{E}^{\pi}[\mu_i^{\pi}(\mathbf{q})]\,\mathbb{E}^{\pi}[w_i^{(s)}(\mathbf{q}))] \\
&= \mathbb{E}^{\pi}\left[\sum_{j=1}^{m} \lambda_j^{\pi}(\mathbf{q}) F_j\left(\lambda_j^{\pi}(\mathbf{q})\right) - \sum_{i=1}^{n} \mu_i^{\pi}(\mathbf{q}) G_i\left(\mu_i^{\pi}(\mathbf{q})\right)\right] - \sum_{j=1}^{m} s_j^{(c)}\,\mathbb{E}^{\pi}[q_j^{(c)}] - \sum_{i=1}^{n} s_i^{(s)}\,\mathbb{E}^{\pi}[q_i^{(s)}].
\end{aligned}
$$

The last step applies Little's Law:

$$
\mathbb{E}[\lambda_j^{\pi}(\mathbf{q})]\,\mathbb{E}^{\pi}[w_j^{(c)}(\mathbf{q})] = \mathbb{E}^{\pi}[q_j^{(c)}] \text{ and } \mathbb{E}^{\pi}[\mu_i^{\pi}(\mathbf{q})]\,\mathbb{E}^{\pi}[w_i^{(s)}(\mathbf{q}))] = \mathbb{E}^{\pi}[q_i^{(s)}].
$$

Note that the last line of the above equation is the long-run average profit under the holding cost model, so the proof is complete. □

C.1.2 Monotonicity of the Optimal Prices (Single-Link Two-Sided Queues)

In this section, we consider a special case with $n = 1$ and $m = 1$, i.e., a single-link two-sided queue given in Figure C.1. The goal of this section is to analyze the optimal pricing policy for this special case, which will motivate our pricing policies for more complex systems.

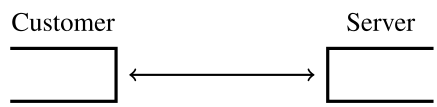


Figure C.1: A single-link two-sided queue.

In single-link systems, there is no incentive for the system operator to hold customers or servers. Whenever possible, we should match the incoming arrival immediately. Thus, at any point of time, there can only be either customers or servers waiting in the system. This enables up to reduce the state space by letting $q = q^{(c)} - q^{(s)}$, the difference between the number of customers and servers waiting in the system. Note that q can be either positive or negative. Using q as the system state, the Bellman equation (4.3) becomes

$$\begin{aligned} h(q) = \max_{\mu \geq 0, \lambda \geq 0} \Bigg[& \frac{F(\lambda)\lambda - G(\mu)\mu}{c} - \frac{s|q|}{c} - \frac{\gamma}{c} + \frac{\lambda}{c} h(q+1) \\ & + \left(1 - \frac{\mu + \lambda}{c}\right) h(q) + \frac{\mu}{c} h(q-1) \Bigg], \forall q \in S \end{aligned} \tag{C.1}$$

where c is a uniformization parameter (see Definition 3) and S denotes the state space. (We omit the subscripts for customer and server types, since $n = m = 1$). We now present the monotonicity result below.

Proposition 12. *For a single-link two-sided queue, there exists an optimal pricing policy* $\mathbf{p}(q) = (p^{(s)}(q), p^{(c)}(q))$, *where both the server price* $p^{(s)}(q)$ *and the customer price* $p^{(c)}(q)$ *increases monotonically with the system state* q.

This result motivates us to search for the optimal pricing policy in the restricted space of monotonic pricing policies, which will be presented in Section C.1.3.

To prove Proposition 12, we first show that the difference of the optimal bias functions, $\Delta h(q) \triangleq h(q) - h(q-1)$, is monotonically decreasing in q. We consider the relative value iteration method to compute $h(q)$. The relative value iteration starts with an arbitrary initial value $h_0(q)$; therefore we choose an initial bias function such that $\Delta h_0(q)$ is decreasing. We also choose any fixed state $q_0 \in S$, say, $q_0 = 0$. In each iteration $k = 1, 2, \cdots$, for all $q \in S$, the relative value iteration algorithm updates the bias function by

$$
\begin{aligned}
h_{k+1}(q) = \max_{\mu \geq 0, \lambda \geq 0} & \Bigg[\frac{F(\lambda)\lambda - G(\mu)\mu}{c} - \frac{s|q|}{c} \\
& + \frac{\lambda}{c} h_k(q+1) + \left(1 - \frac{\mu + \lambda}{c}\right) h_k(q) + \frac{\mu}{c} h_k(q-1) \Bigg] \\
- \max_{\mu \geq 0, \lambda \geq 0} & \Bigg[\frac{F(\lambda)\lambda - G(\mu)\mu}{c} - \frac{s|q_0|}{c} \\
& + \frac{\lambda}{c} h_k(q_0+1) + \left(1 - \frac{\mu + \lambda}{c}\right) h_k(q_0) + \frac{\mu}{c} h_k(q_0-1) \Bigg].
\end{aligned}
$$

As $k \to \infty$, we have $h_k(q) \to h(q)$. Now we will first present a lemma which is essential to prove Proposition 12.

Lemma 16. *The difference of the optimal bias function $\Delta h^*(q)$ that solves* (C.1) *is monotonically decreasing in q.*

Proof of Lemma 16. The proof is by induction using relative value iteration. In the base case, as we can choose any initial bias function in the relative value iteration algorithm, we pick a bias function such that $\Delta h_0(q)$ is decreasing in q, say, $h_0(q) = -q$.

Suppose we are at iteration k. Assume that $\Delta h_k(q)$ is monotonically decreasing for all q. We will now calculate $\Delta h_{k+1}(q+1) - \Delta h_{k+1}(q)$ and show that it is non-positive. The relative value iteration step can be rewritten as

$$
\begin{aligned}
\Delta h_{k+1}(q) = & \max_{\mu \geq 0, \lambda \geq 0} \left[F(\lambda)\lambda - G(\mu)\mu - s|q| + \lambda \Delta h_k(q+1) - \mu \Delta h_k(q) \right] / c \qquad \text{(C.2)} \\
& - \max_{\mu \geq 0, \lambda \geq 0} \left[F(\lambda)\lambda - G(\mu)\mu - s|q-1| + \lambda \Delta h_k(q) - \mu \Delta h_k(q-1) \right] / c + \Delta h_k(q).
\end{aligned}
$$

Thus, we have

$$
\begin{aligned}
&\Delta h_{k+1}(q+1)-\Delta h_{k+1}(q) \\
&\quad = \Delta h_k(q+1)-\Delta h_k(q) \\
&\qquad + \left[\lambda^*(q+1)\Delta h_k(q+2)-\mu^*(q+1)\Delta h_k(q+1)+\mathscr{R}(\mu^*(q+1),\lambda^*(q+1))\right]/c \\
&\qquad - 2\left[\lambda^*(q)\Delta h_k(q+1)-\mu^*(q)\Delta h_k(q)+\mathscr{R}(\mu^*(q),\lambda^*(q))\right]/c \\
&\qquad + \left[\lambda^*(q-1)\Delta h_k(q)-\mu^*(q-1)\Delta h_k(q-1)+\mathscr{R}(\mu^*(q-1),\lambda^*(q-1))\right]/c, \quad \text{(C.3)}
\end{aligned}
$$

where $(\lambda^*(q_0),\mu^*(q_0))$ maximizes the Bellman equation (C.2) and $\mathscr{R}(\mu^*(q),\lambda^*(q))$ is equal to $F(\lambda^*(q))\lambda^*(q)-G(\mu^*(q))\mu^*(q)-s|q|$. Because $(\lambda^*(q),\mu^*(q))$ is the maximizer for state q, we have

$$
\begin{aligned}
&\mathscr{R}(\mu^*(q),\lambda^*(q))+\lambda^*(q)\Delta h_k(q+1)-\mu^*(q)\Delta h_k(q) \\
&\quad \geq \mathscr{R}(\mu^*(q+i),\lambda^*(q+i))+\lambda^*(q+i)\Delta h_k(q+1)-\mu^*(q+i)\Delta h_k(q) \quad \text{for } i\in\{-1,1\}.
\end{aligned}
\quad \text{(C.4)}
$$

Plugging (C.4) to (C.3), we get

$$
\begin{aligned}
\Delta h_{k+1}(q+1)-\Delta h_{k+1}(q) \leq \lambda^*(q+1)\big(\Delta h_k(q+2)-\Delta h_k(q+1)\big)/c + (1-\mu^*(q+1)/c \\
-\lambda^*(q-1)/c)\big(\Delta h_k(q+1)-\Delta h_k(q)\big) \\
+\mu^*(q-1)\big(\Delta h_k(q)-\Delta h_k(q-1)\big)/c.
\end{aligned}
$$

Since the uniformization constant is chosen as $c \geq \lambda_{\max}+\mu_{\max}$, we have $1-\mu^*(q+1)/c-\lambda^*(q-1)/c \geq 0$ for all q. By the induction hypothesis in iteration k, $\Delta h_k(q+1)-\Delta h_k(q) \leq 0$ for all q. Thus, the last line of the above inequality is nonpositive.

This proves that the relative value iteration step preserves the monotonicity of $\Delta h_k(q)$. As $k\to\infty$, $\Delta h_k(q)$ converges to $\Delta h^*(q)$. Since the limit of a decreasing function is decreasing, $\Delta h^*(q)$ is monotonically decreasing. $\square$

Proof of Proposition 12. The equation to be maximize in (C.1) is separable in λ and μ. The domain of λ is $[0, \lambda_{\max}]$ and the domain of μ is $[0, \mu_{\max}]$. By Assumption 2, it is also concave in λ and μ. Thus, (λ^*, μ^*) is an optimal solution to (C.1) either if it is on the boundary of the domain or if it satisfies the first order necessary condition.

First, we consider the boundary case. We will show that if $\lambda^*(q_0) = 0$ for some q_0, then $\lambda^*(q) = 0$ for all $q > q_0$. Suppose $\lambda^*(q_0) = 0$. Then we have

$$F(\lambda)\lambda + \Delta h^*(q_0)\lambda \ \leq \ 0, \quad \forall \lambda \in [0, \lambda_{\max}].$$

By Lemma 16, $\Delta h^*(q)$ is decreasing in q, so

$$\begin{aligned}
F(\lambda)\lambda + \Delta h^*(q_0+k)\lambda \ &= \ F(\lambda)\lambda + \Delta h^*(q_0)\lambda + \left(\Delta h^*(q_0+k) - \Delta h^*(q_0)\right)\lambda \\
&\leq \ \left(\Delta h^*(q_0+k) - \Delta h^*(q_0)\right)\lambda \\
&\leq \ 0, \quad \forall k \geq 1.
\end{aligned}$$

The above inequality implies that $\lambda^*(q_0+k) = 0$. Similarly, if $\lambda^*(q_0) = \lambda_{\max}$ then $\lambda^*(q) = \lambda_{\max}$ for all $q < q_0$. To see this, suppose

$$F(\lambda)\lambda + \Delta h^*(q_0)\lambda \leq F(\lambda_{\max})\lambda_{\max} + \Delta h^*(q_0)\lambda_{\max}, \quad \forall \lambda \in [0, \lambda_{\max}].$$

For any $k \geq 1$, adding $\left(\Delta h^*(q_0-k) - \Delta h^*(q_0)\right)\lambda$ on both sides of the above inequality, we get

$$\begin{aligned}
F(\lambda)\lambda + \Delta h(q_0-k)\lambda &\leq F(\lambda_{\max})\lambda_{\max} + \Delta h^*(q_0)\lambda_{\max} + \left(\Delta h^*(q_0-k) - \Delta h^*(q_0)\right)\lambda \\
&\leq F(\lambda_{\max})\lambda_{\max} + \Delta h^*(q_0)\lambda_{\max} + \left(\Delta h^*(q_0-k) - \Delta h^*(q_0)\right)\lambda_{\max} \\
&= F(\lambda_{\max})\lambda_{\max} + \Delta h^*(q_0-k)\lambda_{\max},
\end{aligned}$$

where the second inequality follows as $\lambda_{\max} \geq \lambda$ and $\Delta h^*(q_0-k) - \Delta h^*(q_0) \geq 0$. Thus,

$\lambda^*(q) = \lambda_{\max}$ is the maximizer for all $q < q_0$. Using the same argument, we can show that if $\mu^*(q_0) = \mu_{\max}$ then $\mu(q)^* = \mu_{\max}$ for all $q \geq q_0$; if $\mu^*(q_0) = 0$ then $\mu^*(q) = 0$ for all $q \leq q_0$.

We now consider the case when the optimal solution in (C.1) is in the interior of the domain, i.e., $\lambda_{\max} > \lambda^*(q) > 0$ and $\mu_{\max} > \mu^*(q) > 0$. By the first-order condition, we have

$$\begin{aligned} [F(\lambda^*(q))\lambda^*(q)]' + \Delta h^*(q+1) = 0, \\ [G(\mu^*(q))\mu^*(q)]' + \Delta h^*(q) = 0. \end{aligned}$$

Note that the derivatives are well defined because F and G are differentiable by Assumption 1. Since $\Delta h^*(q)$ is monotonically decreasing in q by Lemma 16, it implies that $[F(\lambda^*(q))\lambda^*(q)]'$ and $[G(\mu^*(q))\mu^*(q)]'$ are monotonically increasing in q. By Assumption 2, $[F(\lambda)\lambda]'$ is decreasing in λ and $[G(\mu)\mu]'$ is increasing in μ. Thus, by the chain rule, $\lambda^*(q)$ is monotonically decreasing in q and $\mu^*(q)$ is monotonically increasing in q when $\lambda_{\max} > \lambda^* > 0$ and $\mu_{\max} > \mu^* > 0$.

Combining the boundary case and the interior case, we prove that $\lambda^*(q)$ is monotonically decreasing in q. As the demand curve is monotonically decreasing by Assumption 1, the optimal customer price $(p^{(c)})^*(q)$ is monotonically increasing. Also, $\mu^*(q)$ is monotonically increasing in q. As the supply curve is monotonically increasing by Assumption 1, the optimal server price $(p^{(s)})^*(q)$ is monotonically increasing. □

C.1.3 LP-based Approximation Algorithm

Throughout this section, we assume the queue length is bounded by some fixed constant. Specifically, we assume that the state space is given by $S = \{\mathbf{q} \in \mathbb{Z}^{m+n} : \mathbf{0} \leq \mathbf{q} \leq \mathbf{1}_{m+n} q_{\max}\}$ for some $q_{\max} < \infty$. The Bellman equation (4.3) for the uniformized DTMDP defined in the

Section 4.2 can be rewritten as a semi-infinite linear program (see e.g. Bertsekas, 2007):

$$\min_{(\gamma,\mathbf{h})} \ \gamma \tag{C.5}$$

$$\text{subject to} \quad \gamma \geq \mathscr{R}(\mathbf{q},\mathbf{z}) + c\,\mathbb{E}[h(\mathbf{q}') \mid \mathbf{q},\mathbf{z}] - ch(\mathbf{q}) \quad \forall \mathbf{q},\mathbf{z} \in S \times Z(\mathbf{q}), \tag{C.6}$$

where $\mathscr{R}(\mathbf{q},\mathbf{z})$ is the expected profit (see Equation (4.2)) and $\mathbb{E}[h(\mathbf{q}') \mid \mathbf{q},\mathbf{z}]$ is the expectation of the bias function after one transition. However, it is difficult to solve this linear program directly, since it has a separate decision variable $h(\mathbf{q})$ for each state $\mathbf{q}$, and the number of constraints in (C.6) is equal to the number of state-action pairs. Due to the curse of dimensionality, the state space will increase exponentially with the customer and server types.

Polynomial Bias Function Approximation.

We consider an approximation of the MDP in the value space. In particular, we approximate the bias function $h(\mathbf{q})$ by a polynomial function:

$$h(\mathbf{q}) \approx \sum_{l=1}^{r} \left(\sum_{i=1}^{n} b_{l_1}^{(s)} (q_i^{(s)})^l + \sum_{j=1}^{m} b_{l_1}^{(c)} (q_j^{(c)})^l \right) \triangleq \sum_{l=1}^{r} \left\langle \mathbf{b}_l, \mathbf{q}^l \right\rangle, \tag{C.7}$$

for some degree $r \in \mathbb{Z}_+$. Here, $\mathbf{b}_l$ is a vector $(b_{l_1}^{(s)}, \ldots, b_{l_n}^{(s)}, b_{l_1}^{(c)}, \ldots, b_{l_m}^{(c)})$ for all $l \in [r]$.

If we apply this approximation to the linear programming formulation (C.5), it in fact leads to an upper bound approximation of the original MDP (see e.g. Adelman, 2007).

Proposition 13. *Suppose $h(\mathbf{q})$ is replaced with $\sum_{l=1}^{r} \left\langle \mathbf{b}_l, \mathbf{q}^l \right\rangle$ in the optimization problem* (C.5)*. The optimal objective value is an upper bound of the optimal average cost of the original MDP.*

Proof of Proposition 13. Rewriting the Bellman equation using the approximation of the bias function gives us the following semi-infinite linear program:

$$
\begin{aligned}
&\min \;\; \gamma \\
&\text{subject to} \;\; \gamma \geq \mathscr{R}(\mathbf{q},\mathbf{z}) + c\,\mathbb{E}\left[\sum_{l=1}^{r}\left\langle \mathbf{b}_l, \mathbf{q}' \right\rangle \mid \mathbf{q},\mathbf{z}\right] - c\sum_{l=1}^{r}\left\langle \mathbf{b}_l, \mathbf{q}^l \right\rangle \quad \forall \mathbf{q},\mathbf{z} \in S \times Z(\mathbf{q}).
\end{aligned}
$$

The decision variables in the above optimization problem are γ and $\mathbf{b}_l$ $\forall l \in [r]$. Let the optimal solution to the above optimization problem be γ^* and $\mathbf{b}_l^*$ $(\forall l \in [r])$. Now define

$$
h(\mathbf{q}) = \sum_{l=1}^{r}\left\langle \mathbf{b}_l^*, \mathbf{q}^l \right\rangle \;\; \forall \mathbf{q} \in S.
$$

Since $h(\mathbf{q})$ and γ^* are a feasible solution to the optimization problem (C.5), the optimal value of (C.5) is less than or equal to γ^*. $\square$

By approximating the bias function by a polynomial function of degree r, the number of variables is reduced from $q_{\max}^{m+n}$ to $(m+n)\times r$. We will later see that this approximation reduces the computational time of the semi-infinite linear program drastically. As the degree of the polynomial r increases, the upper bound becomes tighter. Therefore, r can be selected to balance the trade-off between approximation accuracy and the computational time.

Matching Policy under Value Function Approximation.

We now focus on matching policies when $h(\mathbf{q})$ is approximated by polynomial functions, in particular, linear and quadratic functions. We denote by $(\boldsymbol{\lambda}^*(\mathbf{q}), \boldsymbol{\mu}^*(\mathbf{q}))$ the optimal pricing decision for state $\mathbf{q}$ under bias function approximation. We can rewrite the Bellman equation as follows:

$$
\frac{\gamma}{c} = \frac{\mathscr{R}(\mathbf{q}, \boldsymbol{\lambda}^*(\mathbf{q}), \boldsymbol{\mu}^*(\mathbf{q}))}{c} + \max_{\mathbf{x} \in X(\mathbf{q})} \left\{ \mathbb{E}\left[h(\mathbf{q}') \mid \mathbf{q},\mathbf{z}\right] - h(\mathbf{q}) \right\}, \quad \forall \mathbf{q} \in S.
$$

The optimal matching decision is given by

$$\mathbf{x}^*(\mathbf{q}) = \underset{\mathbf{x}\in X(\mathbf{q})}{\arg\max} \left\{ \mathbb{E}\left[h(\mathbf{q}') \mid \mathbf{q},\mathbf{z}\right] - h(\mathbf{q}) \right\} \quad \forall \mathbf{q} \in S. \tag{C.8}$$

Suppose we approximate the bias functions by linear functions, say, $h(\mathbf{q}) = \langle -\mathbf{b}, \mathbf{q} \rangle$ for some $\mathbf{b} \in \mathbb{R}^{m+n}$, then we have

$$\begin{aligned}
\mathbf{x}^*(\mathbf{q}) &= \underset{\mathbf{x}\in X(\mathbf{q})}{\arg\max} \left\{ \mathbb{E}\left[h(\mathbf{q}') \mid \mathbf{q},\mathbf{z}\right] - h(\mathbf{q}) \right\} \\
&= \underset{\mathbf{x}\in X(\mathbf{q})}{\arg\max} \left\{ \sum_{j=1}^{m} \lambda_j^*(\mathbf{q}) \left(h(\mathbf{q} + \mathbf{e}_j^{(c)} - \mathbf{x}) - h(\mathbf{q}) \right) \right. \\
&\qquad\qquad \left. + \sum_{i=1}^{n} \mu_i^*(\mathbf{q}) \left(h(\mathbf{q} + \mathbf{e}_i^{(s)} - \mathbf{x}) - h(\mathbf{q}) \right) \right\} \\
&= \underset{\mathbf{x}\in X(\mathbf{q})}{\arg\max} \left\{ \left(\sum_{j=1}^{m} \lambda_j^*(\mathbf{q}) + \sum_{i=1}^{n} \mu_i^*(\mathbf{q}) \right) \langle \mathbf{b}, \mathbf{x} \rangle \right\} \\
&= \underset{\mathbf{x}\in X(\mathbf{q})}{\arg\max} \left\{ \langle \mathbf{b}, \mathbf{x} \rangle \right\}.
\end{aligned}$$

Note that we ignore terms that are independent of $\mathbf{x}$ in the second to last equality. Thus, the optimal matching decision is given by a maximum-weight bipartite matching problem with fixed (vertex) weights $\mathbf{b}$ and feasibility set $X(\mathbf{q})$.

Alternatively, suppose we approximate the bias function by a *quadratic* function of queue lengths given by $h(\mathbf{q}) = \left\langle -\mathbf{1}, \mathbf{q}^2 \right\rangle$. By Equation (C.8), the optimal matching decision is given by

$$
\begin{aligned}
\mathbf{x}^*(\mathbf{q}) =& \arg\max_{\mathbf{x}\in X(\mathbf{q})} \left\{\mathbb{E}\left[h(\mathbf{q}') \mid \mathbf{q},\mathbf{z}\right] - h(\mathbf{q})\right\} \\
=& \arg\max_{\mathbf{x}\in X(\mathbf{q})} \left\{ \sum_{j=1}^{m} \lambda_j^*(\mathbf{q}) \left(h(\mathbf{q}+\mathbf{e}_j^{(c)}-\mathbf{x}) - h(\mathbf{q})\right) \right. \\
& \left. \qquad + \sum_{i=1}^{n} \mu_i^*(\mathbf{q}) \left(h(\mathbf{q}+\mathbf{e}_i^{(s)}-\mathbf{x}) - h(\mathbf{q})\right) \right\} \\
=& \arg\max_{\mathbf{x}\in X(\mathbf{q})} \left\{ \left(\sum_{j=1}^{m} \lambda_j^*(\mathbf{q}) + \sum_{i=1}^{n} \mu_i^*(\mathbf{q})\right) \left(2\langle \mathbf{q},\mathbf{x}\rangle - \left\langle \mathbf{1},\mathbf{x}^2\right\rangle\right) \right. \\
& \left. \qquad + \sum_{j=1}^{m} 2\lambda_j^*(\mathbf{q}) x_j^{(c)} + \sum_{i=1}^{n} 2\mu_i^*(\mathbf{q}) x_i^{(s)} \right\}.
\end{aligned}
$$

We consider two cases. If $X(\mathbf{q}) = \{\mathbf{0}\}$, namely, there is no feasible matching given state $\mathbf{q}$, then we trivially have $\mathbf{x}^*(\mathbf{q}) = \mathbf{0}$. Otherwise, if $X(\mathbf{q}) \neq \{\mathbf{0}\}$, then there exists a matching decision $\mathbf{x} \in X(\mathbf{q})$ such that $2\langle \mathbf{q},\mathbf{x}\rangle - \left\langle \mathbf{1},\mathbf{x}^2\right\rangle > 0$, which means that $\mathbf{x} = \mathbf{0}$ cannot be a maximizer of the above equation. This means that under quadratic value function approximation, customers and servers should be matched instantly when they arrive. Recall that there is at most one new arrival per period in the unformized DTMDP. So, we either have $\left\langle \mathbf{1}, (\mathbf{x}^*(\mathbf{q}))^2\right\rangle = 0$ when $X(\mathbf{q}) = \{\mathbf{0}\}$ or $\left\langle \mathbf{1}, (\mathbf{x}^*(\mathbf{q}))^2\right\rangle = 1+1 = 2$ when $X(\mathbf{q}) \neq \{\mathbf{0}\}$. In either case, we can remove the term $-\left\langle \mathbf{1},\mathbf{x}^2\right\rangle$ in the above equation and get

$$
\begin{aligned}
\mathbf{x}^*(\mathbf{q}) =& \arg\max_{\mathbf{x}\in X(\mathbf{q})} \left\{ \left(\sum_{j=1}^{m} \lambda_j^*(\mathbf{q}) + \sum_{i=1}^{n} \mu_i^*(\mathbf{q})\right) \left(2\langle \mathbf{q},\mathbf{x}\rangle\right) \right. \\
& \left. \qquad + \sum_{j=1}^{m} 2\lambda_j^*(\mathbf{q}) x_j^{(c)} + \sum_{i=1}^{n} 2\mu_i^*(\mathbf{q}) x_i^{(s)} \right\} \qquad \text{(C.9)} \\
\approx& \arg\max_{\mathbf{x}\in X(\mathbf{q})} \left\{ \left(\sum_{j=1}^{m} \lambda_j^*(\mathbf{q}) + \sum_{i=1}^{n} \mu_i^*(\mathbf{q})\right) \left(2\langle \mathbf{q},\mathbf{x}\rangle\right) \right\} \\
=& \arg\max_{\mathbf{x}\in X(\mathbf{q})} \left\{\langle \mathbf{q},\mathbf{x}\rangle\right\},
\end{aligned}
$$

where the approximation step is motivated by the fact that the value of $(\sum_{j=1}^{m} \lambda_j^*(\mathbf{q}) +$

$\sum_{i=1}^{n} \mu_i^*(\mathbf{q}))$ dominates either $\lambda_j^*(\mathbf{q})$ or $\mu_i^*(\mathbf{q})$, especially when n and m are large. This approximation leads to the max-weight matching policy defined in Section 4.2.3.

Constraint Generation Algorithm.

In this section, we use the constraint generation technique to solve the optimization problem (C.5) given approximated bias functions. The constraint generation steps are described in Algorithm 13.

Algorithm 13 Constraint Generation with Bias Function Approximation

---

initialization: $\mathbf{b}_l^0 = 0,\ \forall l \in [r], \gamma^0 = -\infty, k = 0, \varepsilon = +\infty$

initialize master LP: $\{\min \gamma, \text{subject to } \varnothing\}$ Master Problem: (LP^0)

while $\varepsilon >$ tolerance **do**

  $T^k(\mathbf{q},\mathbf{z}) \overset{\Delta}{=} \mathscr{R}(\mathbf{q},\mathbf{z}) + c\,\mathbb{E}\left[\sum_{l=1}^{r} \left\langle \mathbf{b}_l^k, (\mathbf{q}^l)' \right\rangle \mid \mathbf{q},\mathbf{z}\right] - c\sum_{l=1}^{r} \left\langle \mathbf{b}_l^k, \mathbf{q}^l \right\rangle$

  $\delta^k(\mathbf{q}) \leftarrow \max_{\mathbf{z}\in Z(\mathbf{q})} T^k(\mathbf{q},\mathbf{z})$ # Sub-Problem

  if $\delta^k(\mathbf{q}) > \gamma^k$ **then**

    $\mathbf{q}^k, \mathbf{z}^k = \arg\max_{\mathbf{q},\mathbf{z}} T^k(\mathbf{q},\mathbf{z})$

    add the constraint $\gamma \geq T^k(\mathbf{q}^k,\mathbf{z}^k)$ to the master LP (LP^k)

  let $(\gamma^{k+1}, \mathbf{b}_l^{k+1}\ \forall l \in [r])$ be the solution to (LP^k) # Master-Problem

  $\varepsilon \leftarrow \gamma^{k+1} - \gamma^k$

  $k \leftarrow k+1$

output: $\gamma \leftarrow \gamma^k, \mathbf{b}_l \leftarrow \mathbf{b}_l^k,\ \forall l \in [r]$

The algorithm starts with some initial values of the weights $\mathbf{b}$, say $\mathbf{b} = \mathbf{0}$. The algorithm also maintains a master LP that approximates (C.5). In each iteration, the algorithm finds a violating constraint by solving the following sub-problem

$$\max_{\mathbf{q}\leq q_{\max}\mathbf{1}_{m+n}, \mathbf{z}\in Z(\mathbf{q})} \mathscr{R}(\mathbf{q},\mathbf{z}) + c\,\mathbb{E}\left[\sum_{l=1}^{r} \left\langle \mathbf{b}_l^k, (\mathbf{q}^l)' \right\rangle \mid \mathbf{q},\mathbf{z}\right] - c\sum_{l=1}^{r} \left\langle \mathbf{b}_l^k, \mathbf{q}^l \right\rangle. \tag{C.10}$$

If the optimal value of the above subproblem is larger than the optimal value of the master-problem, then a violating constraint is found and added to the master-problem. We then solve the master-problem to get the updated values of $\mathbf{b}_l\ \forall l \in [r]$ and γ. This process is repeated until either no violating constraint is found or the improvement is less than some toleration.

Simulation Results.

We present some simulation results obtained by approximating the bias function by polynomial functions. Figure C.2 shows the pricing policy obtained by linear and quadratic approximation of the bias function, as well as the optimal pricing policy. We observe that both linear and quadratic approximation result in pricing policies that are monotonic, which is consistent with the monotonic structure of the optimal pricing policy. As the penalty coefficient s increases, the approximation error becomes larger.

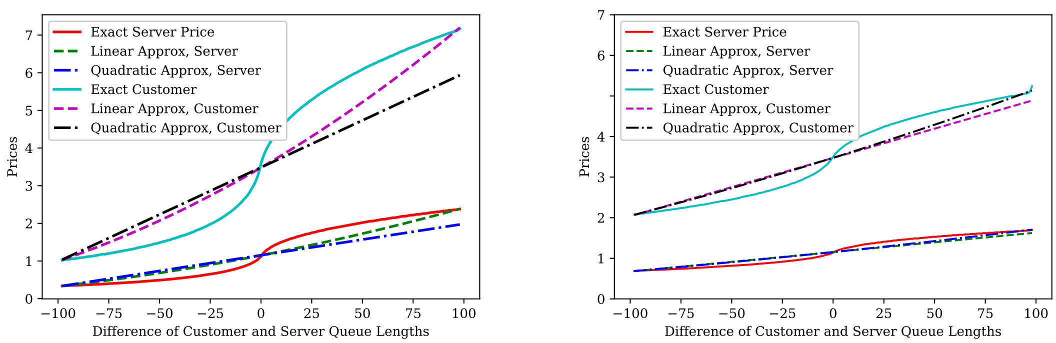


Figure C.2: Comparison of pricing policies with bias approximation $s = 0.05$ (left) and $s = 0.01$ (right).

In the linear approximation case, we also compare the solution from the constraint generation algorithm to the least squares fit of the exact bias function. Our experiment uses both constant elasticity supply/demand functions and linear supply/demand function. The result is presented in Tables C.1 and C.2. We observe again that the approximation error increases with s. We can also see that the optimal objective value γ obtained by constraint generation is an upper bound on the optimal value of the MDP. This verifies Proposition 13.

C.2 Asymptotic Optimality of the Fluid Pricing Policy

C.2.1 Proof of Proposition 5

First note that, under a given pricing and matching policy, if $\mathbb{E}[q_j^{(c)}] = +\infty$ for some $j \in M$ or $\mathbb{E}[q_i^{(s)}] = +\infty$ for some $i \in N$, then $\mathscr{R}(\mathbf{q}, \mathbf{z}) = -\infty$ and the theorem is trivially true, as

Table C.1: Comparison of constraint generation solution with the optimal solution with constant elasticity supply and demand curves.

s	Cons. Generation			Least Squares Fit			% Error	
	b_0	b_1	γ	b_0	b_1	γ	b_0	b_1
0.01	-1.73	-0.007	3.07	-1.69	-0.008	3.03	2%	10%
0.02	-1.73	-0.010	3.07	-1.80	-0.013	3.03	4%	25%
0.05	-1.73	-0.013	3.06	-1.89	-0.019	3.03	9%	34%
0.1	-1.73	-0.014	3.06	-2.05	-0.025	2.95	16%	43%
0.2	-1.73	-0.016	3.06	-2.59	-0.034	2.8	24%	53%
0.5	-1.73	-0.017	3.06	-2.59	-0.051	2.48	33%	67%

Table C.2: Comparison of constraint generation solution with the solution with linear supply and demand curves.

s	Cons. Generation			Least Squares Fit			% Error	
	b_0	b_1	γ	b_0	b_1	γ	b_0	b_1
0.01	-0.014	-2.49	3.11	-0.016	-2.51	3.06	12%	1%
0.02	-0.020	-2.49	3.10	-0.023	-2.51	3.02	13%	1%
0.05	-0.032	-2.48	3.09	-0.037	-2.52	2.93	14%	1%
0.1	-0.048	-2.48	3.07	-0.054	-2.53	2.81	12%	2%
0.2	-0.071	-2.46	3.04	-0.081	-2.55	2.63	12%	3%
0.5	-0.121	-2.44	2.98	-0.144	-2.58	2.24	16%	5%

the optimal objective function value (4.6a) is greater than or equal to 0 because $\tilde{\boldsymbol{\lambda}} = \mathbf{0}_m$ and $\tilde{\boldsymbol{\mu}} = \mathbf{0}_n$ is a feasible solution of the fluid optimization problem. So without loss of generality, we assume that all queue lengths have finite expectations.

We will now show the following lemma.

Lemma 17. *For any stationary pricing and matching policy under which the system is stable and* $\mathbb{E}[q_i^{(s)}] < \infty, \mathbb{E}[q_j^{(c)}] < \infty$ *for all* $i \in N, j \in M$*, the expectation of actions* $\mathbb{E}[\boldsymbol{\lambda}(\mathbf{q})]$*,* $\mathbb{E}[\boldsymbol{\mu}(\mathbf{q})]$ *and* $c \cdot \mathbb{E}[\mathbf{y}(\mathbf{q})]$ *satisfy the constraints in the fluid LP* (4.6b)–(4.6d).

Proof of Lemma 17. We consider the uniformized DTMC induced by a pricing and matching policy such that the system is stable. By the stability condition, for any period k, we have

$$\mathbb{E}[\mathbf{q}(k+1)] = \mathbb{E}[\mathbf{q}(k)],$$

where the expectation is with respect to the stationary distribution of the uniformized DTMC. Let $\mathbf{a}(k)$ be the new arrival at period k and let $\mathbf{x}(k)$ be the matching decision at period k. (Recall that $\mathbf{y}(k) \in \mathbb{Z}_+^{nm}$ represents the matching decision for each customer-server pair and $\mathbf{x}(k) \in \mathbb{Z}_+^{n+m}$ represents the number of matches for each customer and server type.) We can rewrite the stability condition as

$$\mathbb{E}[\mathbf{a}(k)] = \mathbb{E}[\mathbf{x}(k)].$$

Thus, we have

$$c\,\mathbb{E}[\mathbf{x}(k)] = c\,\mathbb{E}[\mathbf{a}(k)] = c\,\mathbb{E}\left[\mathbb{E}[\mathbf{a}(k) \mid \mathbf{q}(k)]\right] = \left(\mathbb{E}[\boldsymbol{\lambda}(\mathbf{q}(k))], \mathbb{E}[\boldsymbol{\mu}(\mathbf{q}(k))]\right),$$

where $\boldsymbol{\lambda}(\mathbf{q}(k))$ and $\boldsymbol{\mu}(\mathbf{q}(k))$ are the arrival rates under the given pricing policy in state $\mathbf{q}(k)$. By Equation (4.1), there exists $y_{ij}(k) \geq 0$ for all $i \in N$ and $j \in M$ such that,

$$\begin{aligned}
x_j^{(c)}(k) &= \sum_{i=1}^{n} y_{ij}(k) \leq q_j^{(c)}(k) \quad \forall j \in M, \\
x_i^{(s)}(k) &= \sum_{j=1}^{m} y_{ij}(k) \leq q_i^{(s)}(k) \quad \forall i \in N, \\
y_{ij}(k) &= 0 \quad \forall (i,j) \notin E.
\end{aligned}$$

Since the matching policy is stationary, the expectation of the matching decision will not depend on k. Taking expectation on both sides with respect to the stationary distribution and defining $\chi_{ij} \triangleq c\,\mathbb{E}[y_{ij}(k)]$, we have

$$\begin{aligned}
\mathbb{E}\left[\lambda_j(k)\right] = c\,\mathbb{E}\left[x_j^{(c)}(k)\right] &= \sum_{i=1}^{n} \chi_{ij}(k) \leq c\,\mathbb{E}\left[q_j^{(c)}(k)\right] < \infty \quad \forall j \in M, \\
\mathbb{E}\left[\mu_i(k)\right] = c\,\mathbb{E}\left[x_i^{(s)}(k)\right] &= \sum_{j=1}^{m} \chi_{ij}(k) \leq c\,\mathbb{E}\left[q_i^{(s)}(k)\right] < \infty \quad \forall i \in N, \\
c\,\mathbb{E}\left[y_{ij}(k)\right] &= \chi_{ij}(k) = 0 \quad \forall (i,j) \notin E.
\end{aligned}$$

Thus, for any pricing and matching policy under which the system is stable, the expectation of actions satisfy the constraints (4.6b)–(4.6d). □

Proof of Proposition 5. Note that the Uniformized DTMC is aperiodic as we will always have transition from a state back to itself. By the ergodic theorem for Markov chains, the long run average profit for a given policy is $\mathbb{E}[\mathscr{R}(\mathbf{q},\mathbf{z})]$. Also, we have

$$\mathbb{E}[\mathscr{R}(\mathbf{q},\mathbf{z})] \leq \mathbb{E}[\langle F(\boldsymbol{\lambda}),\boldsymbol{\lambda}\rangle - \langle G(\boldsymbol{\mu}),\boldsymbol{\mu}\rangle] \leq \langle F(\mathbb{E}[\boldsymbol{\lambda}]),\mathbb{E}[\boldsymbol{\lambda}]\rangle - \langle G(\mathbb{E}[\boldsymbol{\mu}]),\mathbb{E}[\boldsymbol{\mu}]\rangle\,,$$

where the first inequality holds by excluding waiting costs, and the second inequality follows from Jensen's Inequality and Assumption 2. Thus, the optimal value of the fluid problem (4.6a) provides an upper bound for the average profit under any stationary pricing and matching policy. By Lemma 17, a given policy should also satisfy the constraints (4.6b)–(4.6d). □

C.2.2 Proof of Theorem 5

Proof of Theorem 5. By Equation (4.8), the profit-loss L^η is bounded by

$$\begin{aligned} L^\eta \leq & \eta\left(\left\langle F(\boldsymbol{\lambda}^*),(\boldsymbol{\lambda}^*\circ\mathbb{E}[\mathbf{I}^{(c)}(q^\eta_{\max})])\right\rangle - \left\langle G(\boldsymbol{\mu}^*),(\boldsymbol{\mu}^*\circ\mathbb{E}[\mathbf{I}^{(s)}(q^\eta_{\max})]\right\rangle\right) + \langle \mathbf{s},\mathbb{E}[\mathbf{q}]\rangle \\ \leq & \eta\left\langle F(\boldsymbol{\lambda}^*),(\boldsymbol{\lambda}^*\circ\mathbb{E}[\mathbf{I}^{(c)}(q^\eta_{\max})])\right\rangle + \langle \mathbf{s},\mathbb{E}[\mathbf{q}]\rangle\,. \end{aligned} \tag{C.11}$$

To bound the first term, we define a function of customer queue lengths $V^{(c)}(\mathbf{q}) = \sum_{j=1}^m (q_j^{(c)})^2$. Consider the unformized DTMC under the fluid pricing and max-weight matching policy. Suppose $\mathbf{q}$ is the system state in any given period *after* the matching decision $\mathbf{x}$ (or equivalently, $\mathbf{y}$) has been taken. Similarly, let $\mathbf{q}'$ be the state in the next period after the matching decision has been taken. We have

$$
\begin{aligned}
&\eta c\left(\mathbb{E}[V^{(c)}(\mathbf{q}') \mid \mathbf{q}] - V^{(c)}(\mathbf{q})\right) \\
&= \sum_{j=1}^{m} \eta\lambda_j^* \left[\sum_{j'\in M/\{j\}} \left(q_{j'}^{(c)}\right)^2 + \left(q_j + \left(1-\sum_{i=1}^{m} y_{ij}\right)\mathbb{1}_{\left\{q_j^{(c)}<q_{\max}^{\eta}\right\}}\right)^2\right] \\
&\quad + \sum_{i=1}^{m} \eta\mu_i^* \left[\sum_{j=1}^{m}\left(q_j^{(c)} - y_{ij}\mathbb{1}_{\{q_i^{(s)}<q_{\max}^{\eta}\}}\right)^2\right] - \left[\sum_{i=1}^{m}\eta\mu_i^* + \sum_{j=1}^{m}\eta\lambda_j^*\right]\sum_{j=1}^{m}\left(q_j^{(c)}\right)^2 \\
&= \sum_{j=1}^{m} \eta\lambda_j^* \left(1-\sum_{i=1}^{n} y_{ij}\right)\left(1+2q_j^{(c)}\right)\mathbb{1}_{\left\{q_j^{(c)}<q_{\max}^{\eta}\right\}} \\
&\quad + \sum_{i=1}^{n}\eta\mu_i^*\sum_{j=1}^{m} y_{ij}\left(1-2q_j^{(c)}\right)\mathbb{1}_{\{q_i^{(s)}<q_{\max}^{\eta}\}} \\
&\le \sum_{i=1}^{n}\eta\mu_i^* + \sum_{j=1}^{m}\eta\lambda_j^* \\
&\quad + 2\left[\sum_{j=1}^{m}\eta\lambda_j^* q_j^{(c)}\left(1-\sum_{i=1}^{n} y_{ij}\right)\mathbb{1}_{\left\{q_j^{(c)}<q_{\max}^{\eta}\right\}} - \sum_{i=1}^{n}\eta\mu_i^*\sum_{j=1}^{m} q_j^{(c)} y_{ij}\right] \qquad \text{(C.12)} \\
&= \sum_{i=1}^{n}\eta\mu_i^* + \sum_{j=1}^{m}\eta\lambda_j^* \\
&\quad + 2\left[\sum_{j=1}^{m}\eta\lambda_j^* q_j^{(c)}\mathbb{1}_{\left\{q_j^{(c)}<q_{\max}^{\eta}\right\}}\mathbb{1}_{\left\{\max_{i':(i',j)\in E} q_{i'}^{(s)}=0\right\}} - \sum_{i=1}^{n}\eta\mu_i^*\left(\max_{j':(i,j')\in E} q_{j'}^{(c)}\right)\right] \\
&= 2\sum_{(i,j)\in E}\eta\chi_{ij}^*\left[1+q_j^{(c)}\mathbb{1}_{\left\{q_j^{(c)}<q_{\max}^{\eta}\right\}}\mathbb{1}_{\left\{\max_{i':(i',j)\in E} q_{i'}^{(s)}=0\right\}} - \max_{j':(i,j')\in E} q_{j'}^{(c)}\right]. \qquad \text{(C.13)}
\end{aligned}
$$

The first equality holds because the next arrival is a customer of type j with probability $\eta\lambda_j^*/(\eta c)$, and is a server of type i with probability $\eta\mu_i^*/(\eta c)$. The second equality follows as $y_{ij}\in\{0,1\}$. The first inequality follows from the fact that $(1-\sum_{i=1}^{n} y_{ij}) \le 1$ for all $j\in M$ and $\sum_{j=1}^{m} y_{ij}\le 1$ for all $i\in N$, because there can be at most one matching in each period under the max-weight policy. The third equality follows from the definition of the max-weight matching policy (4.5). The last equality follows because $(\boldsymbol{\lambda}^*,\boldsymbol{\mu}^*,\boldsymbol{\chi}^*)$ is the optimal solution to the fluid model and satisfies (4.6b)–(4.6d).

We take expectation with respect to the steady state distribution of $\mathbf{q}$. As $\mathbf{q}\le \mathbf{1}q_{\max}^{\eta}$, $V^{(c)}(\mathbf{q})$ is bounded, so $\mathbb{E}[V^{(c)}(\mathbf{q})]$ is finite. By the assumption of that $\mathbf{q}$ follows the steady state distribution, we have $\mathbf{E}[V^{(c)}(\mathbf{q}') - V^{(c)}(\mathbf{q})] = 0$, so the expectation of the left hand side is 0. Therefore,

$$\begin{aligned}
0 &\le \sum_{(i,j)\in E} \chi^*_{ij} \left[1 + \mathbb{E}\left[q_j^{(c)} \mathbb{1}_{\left\{q_j^{(c)} < q^\eta_{\max}\right\}} \mathbb{1}_{\left\{\max_{i':(i',j)\in E} q_{i'}^{(s)} = 0\right\}} \right] - \mathbb{E}\left[\max_{j':(i,j')\in E} q_{j'}^{(c)} \right] \right] \\
&\le \sum_{(i,j)\in E} \chi^*_{ij} \left[1 + \mathbb{E}\left[q_j^{(c)} \right] - \mathbb{E}\left[I_j^{(c)}(q^\eta_{\max}) \right] q^\eta_{\max} - \mathbb{E}\left[\max_{j':(i,j')\in E} q_{j'}^{(c)} \right] \right] \\
&\le \sum_{(i,j)\in E} \chi^*_{ij} \left[1 - \left[I_j^{(c)}(q^\eta_{\max}) \right] q^\eta_{\max} \right] \\
&= \sum_{j=1}^m \lambda^*_j - q^\eta_{\max} \sum_{j=1}^m \lambda^*_j \left[I_j^{(c)}(q^\eta_{\max}) \right]. \qquad \text{(C.14)}
\end{aligned}$$

The second inequality above holds because $\mathbb{1}_{\max_{i'\in N(j)} q_{i'}^{(s)} = 0} \le 1$ and $q_j^{(c)} \le q^\eta_{\max}$. (Recall that $I_j^{(c)}(q^\eta_{\max}) \triangleq \mathbb{1}_{\left\{q_j^{(c)} = q^\eta_{\max}\right\}}$.) The third inequality holds because $\max_{j'\in N(i)} q_{j'}^{(c)} \ge q_j^{(c)}$ for all j such that $(i,j)\in E$. By substituting $q^\eta_{\max} = \gamma\sqrt{\eta}$ in (C.14) for an arbitrary positive constant γ, we get

$$\gamma\sqrt{\eta} \sum_{j=1}^m \lambda^*_j \mathbb{E}\left[I_j^{(c)}(q^\eta_{\max}) \right] \le \sum_{j=1}^m \lambda^*_j,$$

which implies

$$\sum_{j=1}^m F_j(\lambda^*_j) \lambda^*_j \mathbb{E}\left[I_j^{(c)}(q^\eta_{\max}) \right] \le \frac{1}{\gamma\sqrt{\eta}} \max_{j\in M} F_j(\lambda^*_j) \sum_{j=1}^m \lambda^*_j.$$

Thus, the first term $\left\langle F(\boldsymbol{\lambda}^*), \boldsymbol{\lambda}^* \circ \mathbb{E}\left[\mathbf{I}^{(c)}(q^\eta_{\max}) \right] \right\rangle$ of (C.11) is $O(1/\sqrt{\eta})$.

Now we will bound the second term in (C.11). The queue length $\mathbf{q}$ under the fluid pricing policy always satisfies $\mathbf{q} \le \mathbf{1}_{n+m} q^\eta_{\max}$. Thus, it is trivially true that $\langle \mathrm{s}, \mathbb{E}[\mathbf{q}] \rangle \le q^\eta_{\max} \langle \mathbf{1}_{n+m}, \mathrm{s} \rangle = \gamma\sqrt{\eta} \langle \mathbf{1}_{n+m}, \mathrm{s} \rangle$. Thus we can upper bound the profit loss L^η by using (4.8) as follows:

$$L^\eta \le \sqrt{\eta} \left(\max_{j\in M} \frac{F_j(\lambda^*_j)}{\gamma} \sum_{j=1}^m \lambda^*_j + \gamma \langle \mathbf{1}_{n+m}, \mathrm{s} \rangle \right) = O(\sqrt{\eta}). \qquad \text{(C.15)}$$

$\square$

C.2.3 Proof of Proposition 6

Proof of Proposition 6. To prove the lower bound of profit loss, we only need to consider single-link queues, that is, $n = m = 1$. In the proof, we omitted the subscript for the type of customers and servers. Under the fluid pricing policy, the steady state distribution of $q = q^{(c)} - q^{(s)}$ is uniform in $[-q^{\eta}_{\max}, q^{\eta}_{\max}]$, as q behaves like a symmetric simple random walk. Thus, the expected value of the sum of queue length $q^{(s)} + q^{(c)}$ can be computed in terms of the buffer capacity $q^{\eta}_{\max}$ as follows:

$$\begin{aligned}\mathbb{E}[q^{(s)}(k) + q^{(c)}(k)] &= \mathbb{E}[|q^{(s)}(k) - q^{(c)}(k)|] \\ &= \frac{q^{\eta}_{\max}(q^{\eta}_{\max} + 1)}{2q^{\eta}_{\max} + 1}.\end{aligned}$$

The probabilities of $q^{(s)} = q^{\eta}_{\max}$ and $q^{(c)} = q^{\eta}_{\max}$ are

$$\begin{aligned}\Pr[q^{(s)} = q^{\eta}_{\max}] &= \frac{1}{2q^{\eta}_{\max} + 1}, \\ \Pr[q^{(c)} = q^{\eta}_{\max}] &= \frac{1}{2q^{\eta}_{\max} + 1}.\end{aligned}$$

By Equation (4.8), the expected profit loss is bounded by

$$\begin{aligned}L^{\eta} &\le \frac{F(\lambda^*)\lambda^* - G(\mu^*)\mu^*}{2q^{\eta}_{\max} + 1}\eta + (s^{(s)} + s^{(c)})\frac{q^{\eta}_{\max}(q^{\eta}_{\max} + 1)}{2q^{\eta}_{\max} + 1}, \\ L^{\eta} &\ge \frac{F(\lambda^*)\lambda^* - G(\mu^*)\mu^*}{2q^{\eta}_{\max} + 1}\eta + \min\{s^{(s)}, s^{(c)}\}\frac{q^{\eta}_{\max}(q^{\eta}_{\max} + 1)}{2q^{\eta}_{\max} + 1}. \qquad \text{(C.16)}\end{aligned}$$

The inequalities above show that the profit loss is minimized with respect to η if $q^{\eta}_{\max} = \gamma\sqrt{\eta}$ for some positive constant γ. To see this, if $q^{\eta}_{\max} = \gamma\eta^{0.5+\varepsilon}$ for some $\varepsilon > 0$, then due to the second term in (C.16), $L^{\eta} = \Theta(\eta^{0.5+\varepsilon})$. On the other hand if $\phi^{\eta}_{\max} = \gamma\eta^{0.5-\varepsilon}$ for some $\varepsilon > 0$, then due to the first term in (C.16), $L^{\eta} = \Theta(\eta^{0.5+\varepsilon})$. Therefore, by taking

$q_{\max}^{\eta} = \gamma\sqrt{\eta}$ in (C.16), the optimal profit loss is $L^{\eta} = \Theta(\sqrt{\eta})$. □

C.3 Asymptotic Optimality of the Two-Price Policy

C.3.1 Proof of Lemma 1

Proof of Lemma 1. We start by defining two functions below:

$$V^{(s)}(\mathbf{q}) \triangleq \left\langle \mathbf{1}_n, (\mathbf{q}^{(s)})^2 \right\rangle, \quad V^{(c)}(\mathbf{q}) \triangleq \left\langle \mathbf{1}_m, (\mathbf{q}^{(c)})^2 \right\rangle.$$

Consider the uniformized DTMC for the η^{th} scaled system under the two-price pricing policy. Suppose $\mathbf{q}$ is the system state *after* the matching decision $\mathbf{x}$ (or equivalently, $\mathbf{y}$) has been taken. We use $\mathbf{q}'$ to denote the state in the next period. The drift of $V^{(c)}(\mathbf{q})$ after one transition is given by

$$
\begin{aligned}
&\eta c\left(\mathbb{E}\left[V^{(c)}(\mathbf{q}') \mid \mathbf{q}\right]-V^{(c)}(\mathbf{q})\right) \\
&=\sum_{j=1}^{m}\left(\eta\lambda_j^*-\theta_j\sigma^\eta\mathbb{1}_{\left\{q_j^{(c)}>\tau_{\max}^\eta\right\}}\right)\left(\left(q_j^{(c)}+1-\sum_{i=1}^{n}y_{ij}\right)^2-\left(q_j^{(c)}\right)^2\right) \\
&\quad+\sum_{i=1}^{n}\left(\eta\mu_i^*-\phi_i\sigma^\eta\mathbb{1}_{\left\{q_i^{(s)}>\tau_{\max}^\eta\right\}}\right)\sum_{j=1}^{m}\left(\left(q_j^{(c)}-y_{ij}\right)^2-\left(q_j^{(c)}\right)^2\right) \\
&\overset{(a)}{=}\sum_{j=1}^{m}\left(\eta\lambda_j^*-\theta_j\sigma^\eta\mathbb{1}_{\left\{q_j^{(c)}>\tau_{\max}^\eta\right\}}\right)\left(1-\sum_{i=1}^{n}y_{ij}\right)\left(1+2q_j^{(c)}\right) \\
&\quad+\sum_{i=1}^{n}\left(\eta\mu_i^*-\phi_i\sigma^\eta\mathbb{1}_{\left\{q_i^{(s)}>\tau_{\max}^\eta\right\}}\right)\sum_{j=1}^{m}y_{ij}\left(1-2q_j^{(c)}\right) \\
&\overset{(b)}{\leq}\eta\left\langle\mathbf{1}_m,\boldsymbol{\lambda}^*\right\rangle+\eta\left\langle\mathbf{1}_n,\boldsymbol{\mu}^*\right\rangle+2\sum_{j=1}^{m}\eta\lambda_j^*\left(1-\sum_{i=1}^{n}y_{ij}\right)q_j^{(c)}-2\sum_{i=1}^{n}\eta\mu_i^*\sum_{j=1}^{m}y_{ij}q_j^{(c)} \\
&\quad-2\sigma^\eta\sum_{j=1}^{m}\theta_j\left(1-\sum_{i=1}^{n}y_{ij}\right)q_j^{(c)}\mathbb{1}_{\left\{q_j^{(c)}>\tau_{\max}^\eta\right\}}+2\sigma^\eta\sum_{i=1}^{n}\phi_i\sum_{j=1}^{m}y_{ij}q_j^{(c)}\mathbb{1}_{\left\{q_i^{(s)}>\tau_{\max}^\eta\right\}} \\
&\overset{(c)}{=}\eta\left\langle\mathbf{1}_m,\boldsymbol{\lambda}^*\right\rangle+\eta\left\langle\mathbf{1}_n,\boldsymbol{\mu}^*\right\rangle+2\sum_{j=1}^{m}\eta\lambda_j^*q_j^{(c)}\mathbb{1}_{\left\{\max_{i':(i',j)\in E}q_{i'}^{(s)}=0\right\}} \\
&\quad-2\sum_{i=1}^{n}\eta\mu_i^*\max_{j':(i,j')\in E}q_{j'}^{(c)}-2\sigma^\eta\sum_{j=1}^{m}\theta_jq_j^{(c)}\mathbb{1}_{\left\{q_j^{(c)}>\tau_{\max}^\eta\right\}}\mathbb{1}_{\left\{\max_{i':(i',j)\in E}q_{i'}^{(s)}=0\right\}} \\
&\quad+2\sigma^\eta\sum_{i=1}^{n}\phi_i\max_{j':(i,j')\in E}q_{j'}^{(c)}\mathbb{1}_{\left\{q_i^{(s)}>\tau_{\max}^\eta\right\}} \qquad \text{(C.17)} \\
&\overset{(d)}{=}\eta\left\langle\mathbf{1}_m,\boldsymbol{\lambda}^*\right\rangle+\eta\left\langle\mathbf{1}_n,\boldsymbol{\mu}^*\right\rangle+2\eta\sum_{i,j\in E}\chi_{ij}^*\left(q_j^{(c)}\mathbb{1}_{\left\{\max_{i':(i',j)\in E}q_{i'}^{(s)}=0\right\}}-\max_{j':(i,j')\in E}q_{j'}^{(c)}\right) \\
&\quad-2\sigma^\eta\sum_{j=1}^{m}\theta_jq_j^{(c)}\mathbb{1}_{\left\{q_j^{(c)}>\tau_{\max}^\eta\right\}} \qquad \text{(C.18)} \\
&\overset{(e)}{\leq}\eta\left\langle\mathbf{1}_m,\boldsymbol{\lambda}^*\right\rangle+\eta\left\langle\mathbf{1}_n,\boldsymbol{\mu}^*\right\rangle-2\sigma^\eta\sum_{j=1}^{m}\theta_jq_j^{(c)}\mathbb{1}_{\left\{q_j^{(c)}>\tau_{\max}^\eta\right\}}. \qquad \text{(C.19)}
\end{aligned}
$$

Under the max-weight matching policy, new arrivals are immediately matched to compatible counterparts if their queues are nonempty. Thus, $\sum_{i=1}^{n}y_{ij}$ and y_{ij} (for all j) are either 1 or 0. Step (a) then follows from the fact that $(1-\sum_{i=1}^{n}y_{ij})^2=1-\sum_{i=1}^{n}y_{ij}$ and $(y_{ij})^2=y_{ij}$. Step (b) holds because $\eta\lambda_j^*-\theta_j\sigma^\eta\mathbb{1}_{\left\{q_j^{(c)}>\tau_{\max}^\eta\right\}}<\eta\lambda_j^*$ as $\theta_j\sigma^\eta>0$ and $1-\sum_{i=1}^{n}y_{ij}\leq 1$, and because $\eta\mu_i^* - \phi_i\sigma^\eta \mathbb{1}_{\left\{q_i^{(s)} > \tau_{\max}^\eta\right\}} < \eta\mu_i^*$ as $\phi_i\sigma^\eta > 0$ and $\sum_{j=1}^m y_{ij} \leq 1$. Step (c) follows from the definition of the max-weight policy in (4.5). In Eq (C.17), the event $q_i^{(s)} > \tau_{\max}^\eta$ implies that all compatible types of i have empty queues. Similarly, the event $q_j^{(c)} > \tau_{\max}^\eta$ implies that all compatible types of j have empty queues. Then, step (d) follows from the definition of the optimal fluid solution $(\boldsymbol{\lambda}^*, \boldsymbol{\mu}^*, \boldsymbol{\xi}^*)$ given by (4.6). Lastly, step (e) follows because $q_j^{(c)} \leq \max_{j':(i,j')\in E} q_{j'}^{(c)}$.

By the same argument, we can bound the drift of $V^{(s)}(\mathbf{q})$ by

$$\begin{aligned}
&\eta c \left(\mathbb{E}[V^{(s)}(\mathbf{q}') \mid \mathbf{q}] - V^{(s)}(\mathbf{q})\right) \\
&\qquad \leq \eta \langle \mathbf{1}_m, \boldsymbol{\lambda}^* \rangle + \eta \langle \mathbf{1}_n, \boldsymbol{\mu}^* \rangle - 2\sigma^\eta \sum_{i=1}^n \phi_i q_i^{(s)} \mathbb{1}_{\left\{q_i^{(s)} > \tau_{\max}^\eta\right\}}. \qquad \text{(C.20)}
\end{aligned}$$

Define $V(\mathbf{q}) \triangleq V^{(s)}(\mathbf{q}) + V^{(c)}(\mathbf{q})$. Adding (C.19) and (C.20), we can bound the drift of $V(\mathbf{q})$ by

$$\eta c \left(\mathbb{E}[V(\mathbf{q}') \mid \mathbf{q}] - V^{(s)}(\mathbf{q})\right) \leq B - 2\sigma^\eta \sum_{i=1}^n \phi_i q_i^{(s)} \mathbb{1}_{\left\{q_i^{(s)} > \tau_{\max}^\eta\right\}} - 2\sigma^\eta \sum_{j=1}^m \theta_j q_j^{(c)} \mathbb{1}_{\left\{q_j^{(c)} > \tau_{\max}^\eta\right\}},$$

where $B \triangleq 2\eta \langle \mathbf{1}_m, \boldsymbol{\lambda}^* \rangle + 2\eta \langle \mathbf{1}_n, \boldsymbol{\mu}^* \rangle > 0$ is a positive constant that is independent of $\mathbf{q}$.

Consider the following finite set:

$$\begin{aligned}
\mathscr{B}^\eta = \Bigg\{ &\mathbf{q} \in \mathbb{Z}_+^{m+n} : \\
&q_i^{(s)} \leq \max\left\{\frac{B}{\phi_i\sigma^\eta}, \tau_{\max}^\eta\right\}, q_j^{(c)} \leq \max\left\{\frac{B}{\theta_j\sigma^\eta}, \tau_{\max}^\eta\right\} \ \forall i \in N, j \in M \Bigg\}.
\end{aligned}$$

Outside the finite set $\mathscr{B}^\eta$, the drift of the Lyapunov function $V(\mathbf{q})$ is strictly negative. Specifically, we have

$$\eta c \left(\mathbb{E}[V(\mathbf{q}') \mid \mathbf{q}] - V^{(s)}(\mathbf{q})\right) \leq -B, \quad \forall \mathbf{q} \notin \mathscr{B}^\eta. \qquad \text{(C.21)}$$

Thus, by the Foster-Lyapunov theorem (Srikant and Ying, 2014), the uniformized DTMC under the two-price and max-weight policy is positive recurrent for any η. The first half of the lemma is proved.

To prove the second half of the lemma, we apply the moment bound theorem (Hajek, 2006) to Equation (C.21), which gives that

$$\sigma^{\eta}\,\mathbb{E}\left[\sum_{i=1}^{n}\phi_i q_i^{(s)}\mathbb{1}_{\left\{q_i^{(s)}>\tau_{\max}^{\eta}\right\}}+\sum_{j=1}^{m}\theta_j q_j^{(c)}\mathbb{1}_{\left\{q_j^{(c)}>\tau_{\max}^{\eta}\right\}}\right]\leq\frac{B}{2},$$

where the expectation is taken over the stationary distribution under the two-price and max-weight policy. By substituting $\mathbb{1}_{\left\{q_i^{(s)}>\tau_{\max}^{\eta}\right\}}=1-\mathbb{1}_{\left\{q_i^{(s)}\leq\tau_{\max}^{\eta}\right\}}$ and then using the fact that $\mathbb{E}[q_i^{(s)}\mathbb{1}_{\left\{q_i^{(s)}\leq\tau_{\max}^{\eta}\right\}}]\leq\tau_{\max}^{\eta}$, we get the desired result. □

C.3.2 Proof of Lemma 2

Before proving Lemma 2, we first prove the following claim.

Claim 1. *Consider a pricing and matching policy that has the following form:*

$$\lambda_j(\mathbf{q})=\eta\lambda_j^*+\tilde{f}_j(\mathbf{q},\eta)\quad\forall j\in M, \tag{C.22}$$

$$\mu_i(\mathbf{q})=\eta\mu_i^*+\tilde{g}_i(\mathbf{q},\eta)\quad\forall i\in N. \tag{C.23}$$

Suppose the system is positive recurrent under this policy. Then we have

$$\begin{aligned}&\sum_{j\in M}\left(F_j'(\lambda_j^*)\lambda_j^*+F_j(\lambda_j^*)\right)\mathbb{E}[\tilde{f}_j(\mathbf{q},\eta)]-\sum_{i\in N}\left(G_i'(\mu_i^*)\mu_i^*+G_i(\mu_i^*)\right)\mathbb{E}[\tilde{g}_i(\mathbf{q},\eta)]\\&\qquad=-\frac{1}{\eta}\sum_{(i,j)\in E}\xi_{ij}\bar{\chi}_{ij}\mathbb{1}_{\chi_{ij}^*=0},\end{aligned}$$

where $\bar{\chi}_{ij}$ and ξ_{ij} are some nonnegative constants.

Proof of Claim 1. Denote the objective function (4.6a) of the fluid model by $r^{\eta}:\mathbb{R}^{m+n}\to\mathbb{R}$. Denote the left-hand sides of the constraints (4.6b)–(4.6c) by $\mathbf{h}:\mathbb{R}^{m+n+mn}\to\mathbb{R}^{m+n}$.

Recall that $(\eta\boldsymbol{\lambda}^*, \eta\boldsymbol{\mu}^*, \eta\boldsymbol{\chi}^*)$ is an optimal solution to the fluid problem. Since the fluid problem is convex and satisfies Slater's condition, by the KKT conditions, there exists Lagrange multipliers $\boldsymbol{\kappa} \in \mathbb{R}^{m+n}$ and $\boldsymbol{\xi} \in \mathbb{R}_+^{m\times n}$ such that

$$\nabla r^\eta(\eta\boldsymbol{\lambda}^*, \eta\boldsymbol{\mu}^*, \eta\boldsymbol{\chi}^*) + \nabla \mathbf{h}(\eta\boldsymbol{\lambda}^*, \eta\boldsymbol{\mu}^*, \eta\boldsymbol{\chi}^*)^\top \boldsymbol{\kappa} + \sum_{(i,j)\in E} \xi_{ij}\mathbf{e}_{ij} + \sum_{(i,j)\notin E} \xi_{ij}\mathbf{e}_{ij} = \mathbf{0}, \tag{C.24}$$

where

$$\nabla r^\eta(\eta\boldsymbol{\lambda}^*, \eta\boldsymbol{\mu}^*, \eta\boldsymbol{\chi}^*) = \big(\eta F'(\boldsymbol{\lambda}^*)\boldsymbol{\lambda}^* + \eta F(\boldsymbol{\lambda}^*),\ -\eta G'(\boldsymbol{\mu}^*)\boldsymbol{\mu}^* - \eta G(\boldsymbol{\mu}^*),\ \mathbf{0}\big)$$

and $\mathbf{e}_{ij} \in \mathbb{R}^{m+n+mn}$ is a vector with the component $m+n+(i,j)$ being 1 and all other components being 0.

Let $\bar{\boldsymbol{\lambda}} = \mathbb{E}[\boldsymbol{\lambda}(\mathbf{q})]$, $\bar{\boldsymbol{\mu}} = \mathbb{E}[\boldsymbol{\mu}(\mathbf{q})]$. Since the system is positive recurrent under the given policy, by Lemma 17, there exists a vector $\bar{\boldsymbol{\chi}}$ such that the constraints (4.6b)–(4.6d) of the fluid model are satisfied. We define a vector $\mathbf{d} \in \mathbb{R}^{n+m+mn}$ given by

$$\mathbf{d} = (\eta\boldsymbol{\lambda}^*, \eta\boldsymbol{\mu}^*, \eta\boldsymbol{\chi}^*) - (\bar{\boldsymbol{\lambda}}, \bar{\boldsymbol{\mu}}, \bar{\boldsymbol{\chi}}) = (-\mathbb{E}[\tilde{\mathbf{f}}(\mathbf{q},\eta)],\ -\mathbb{E}[\tilde{\mathbf{g}}(\mathbf{q},\eta)],\ \eta\boldsymbol{\chi}^* - \bar{\boldsymbol{\chi}}).$$

Because the function $\mathbf{h}$ is linear, it holds that

$$\langle \nabla \mathbf{h}, \mathbf{d}\rangle = \mathbf{h}(\eta\boldsymbol{\lambda}^*, \eta\boldsymbol{\mu}^*, \eta\boldsymbol{\chi}^*) - \mathbf{h}(\bar{\boldsymbol{\lambda}}, \bar{\boldsymbol{\mu}}, \bar{\boldsymbol{\chi}}) = \mathbf{0}. \tag{C.25}$$

By complementary slackness, we have

$$\xi_{ij}(\eta\chi_{ij}^* - \bar{\chi}_{ij}) = -\xi_{ij}\bar{\chi}_{ij}\mathbb{1}_{\chi_{ij}^*=0}, \quad \forall (i,j)\in E. \tag{C.26}$$

Moreover, since $\chi^*_{ij} = 0$ and $\bar{\chi}_{ij} = 0$ for all $(i, j) \notin E$, we have

$$\xi_{ij}(\eta\chi^*_{ij} - \bar{\chi}_{ij}) = 0, \quad \forall (i, j) \notin E. \tag{C.27}$$

Thus, by taking the inner product on both the side by $\mathbf{d}$ in (C.24) and using (C.25), (C.26) and (C.27), we get the desired result. □

Proof of Lemma 2. According to definition of the two-price policy (4.9), for all $j \in M$, we have

$$\begin{aligned}
\mathbb{E}[\lambda^{(c)}_j(\mathbf{q})] &= \mathbb{E}\left[\lambda^{(c)}_j(\mathbf{q}) \mid q^{(c)}_j \leq \tau^{\eta}_{\max}\right] \Pr[q^{(c)}_j \leq \tau^{\eta}_{\max}] \\
&\quad + \mathbb{E}\left[\lambda^{(c)}_j(\mathbf{q}) \mid q^{(c)}_j > \tau^{\eta}_{\max}\right] \Pr[q^{(c)}_j > \tau^{\eta}_{\max}] \\
&= \eta\lambda^*_j - \theta_j\sigma^{\eta} \Pr[q^{(c)}_j > \tau^{\eta}_{\max}].
\end{aligned}$$

Similarly, for all $i \in N$, we have

$$\mathbb{E}[\mu^{(s)}_i(\mathbf{q})] = \eta\mu^*_i - \phi_i\sigma^{\eta} \Pr[q^{(s)}_i > \tau^{\eta}_{\max}].$$

Define a vector $\mathbf{d} \in \mathbb{R}^{n+m+mn}$ given by

$$\mathbf{d} = (\eta\boldsymbol{\lambda}^*, \eta\boldsymbol{\mu}^*, \eta\bar{\boldsymbol{\chi}}^*) - (\mathbb{E}[\boldsymbol{\lambda}(\mathbf{q})], \mathbb{E}[\boldsymbol{\mu}(\mathbf{q})], \mathbb{E}[\mathbf{y}(\mathbf{q})]),$$

where $\mathbf{y}(\mathbf{q}) \in \mathbb{R}^{mn}_+$ is the matching decision under the two-price policy in state $\mathbf{q}$. We let $\chi^{TP}_{ij} = \mathbb{E}[y_{ij}(\mathbf{q})]$.

By Lemma 1, the system is positive recurrent under the two-price policy for any $\boldsymbol{\theta}^{\eta} > \mathbf{0}_m$, $\boldsymbol{\phi}^{\eta} > \mathbf{0}_n$ and $\boldsymbol{\sigma}^{\eta}$ with $\mathbb{E}[\langle \mathbf{1}_{m+n}, \mathbf{q}\rangle] < \infty$. Also, the two-price policy falls under the form given in (C.22) and (C.23). Thus, by Claim 1, we have

$$
\begin{aligned}
&\sum_{j\in M}\left(F_j'(\lambda_j^*)\lambda_j^*+F_j(\lambda_j^*)\right)\theta_j \Pr[q_j^{(c)}>\tau_{\max}^\eta]-\sum_{i\in N}\left(G_i'(\mu_i^*)\mu_i^*+G_i(\mu_i^*)\right)\phi_i \Pr[q_i^{(s)}>\tau_{\max}^\eta]\\
&\qquad=\frac{1}{\eta}\sum_{(i,j)\in E}\xi_{ij}\chi_{ij}^{TP}\mathbb{1}_{\chi_{ij}^*=0}\\
&\qquad=\sum_{(i,j)\in E}\xi_{ij}\left|\chi_{ij}^*-\frac{1}{\eta}\chi_{ij}^{TP}\right|\\
&\qquad\le |E|^2\max_{i\in N,j\in M}\{\theta_j,\phi_i\}\frac{\sigma^\eta}{\eta},
\end{aligned}
$$

where the last inequality follows from Theorem 10.5 of Schrijver (1998). □

C.3.3 Proof of Theorem 6

Proof of Theorem 6. We will first calculate the profit loss given by (4.8) as follows:

$$
\begin{aligned}
L^\eta &= \gamma_*^\eta-(\gamma^\eta-\langle \mathbf{s},\mathbb{E}[\mathbf{q}]\rangle)\\
&=\eta\,\langle F(\boldsymbol{\lambda}^*),\boldsymbol{\lambda}^*\rangle-\eta\,\langle G(\boldsymbol{\mu}^*),\boldsymbol{\mu}^*\rangle-\eta\left(\sum_{j\in M}F_j(\lambda_j^*)\lambda_j^*\Pr[q_j^{(c)}\le\tau_{\max}^\eta]+\right.\\
&\quad\sum_{j\in M}F_j\left(\lambda_j^*-\frac{\theta_j\sigma^\eta}{\eta}\right)\left(\lambda_j^*-\frac{\theta_j\sigma^\eta}{\eta}\right)\Pr[q_j^{(c)}>\tau_{\max}^\eta]-\sum_{i\in N}G_i(\mu_i^*)\mu_i^*\Pr[q_i^{(s)}\le\tau_{\max}^\eta]\\
&\quad\left.-\sum_{i\in N}G_i\left(\mu_i^*-\frac{\phi_i\sigma^\eta}{\eta}\right)\left(\mu_i^*-\frac{\phi_i\sigma^\eta}{\eta}\right)\Pr[q_i^{(s)}>\tau_{\max}^\eta]\right)+\langle \mathbf{s},\mathbb{E}[\mathbf{q}]\rangle.\\
&=\eta\sum_{j\in M}\left(F_j(\lambda_j^*)\lambda_j^*-F_j\left(\lambda_j^*-\frac{\theta_j\sigma^\eta}{\eta}\right)\left(\lambda_j^*-\frac{\theta_j\sigma^\eta}{\eta}\right)\right)\Pr[q_j^{(c)}>\tau_{\max}^\eta]\\
&\quad-\eta\sum_{i\in N}\left(G_i(\mu_i^*)\mu_i^*-G_i\left(\mu_i^*-\frac{\phi_i\sigma^\eta}{\eta}\right)\left(\mu_i^*-\frac{\phi_i\sigma^\eta}{\eta}\right)\right)\Pr[q_i^{(s)}>\tau_{\max}^\eta]+\langle \mathbf{s},\mathbb{E}[\mathbf{q}]\rangle.
\end{aligned}
\tag{C.28}
$$

We apply Taylor's theorem to the terms $F_j(\lambda_j^* - \theta_j\sigma^\eta/\eta)$ and $G_i(\mu_i^* - \phi_i\sigma^\eta/\eta)$. For customer type j, using $\sigma^\eta = \eta^{2/3}$, we have

$$\begin{aligned}
&\eta F_j(\lambda_j^*)\lambda_j^* - \eta F_j\left(\lambda_j^* - \frac{\theta_j\sigma^\eta}{\eta}\right)\left(\lambda_j^* - \frac{\theta_j\sigma^\eta}{\eta}\right) \\
&= \eta F_j(\lambda_j^*)\lambda_j^* \\
&\quad - \eta\left(F_j(\lambda_j^*) - \frac{\theta_j\sigma^\eta}{\eta}F_j'(\lambda_j^*) + \left(\frac{\theta_j\sigma^\eta}{\eta}\right)^2 F_j''(\lambda_j^*) + o(\eta^{-2/3})\right)\left(\lambda_j^* - \frac{\theta_j\sigma^\eta}{\eta}\right) \\
&= \left(F_j(\lambda_j^*) + F_j'(\lambda_j^*)\lambda_j^*\right)\theta_j\sigma^\eta - \left(F_j''(\lambda_j^*)\lambda_j^* + F_j'(\lambda_j^*)\right)\frac{(\theta_j\sigma^\eta)^2}{\eta} + o(\eta^{1/3}) \\
&= \left(F_j(\lambda_j^*) + F_j'(\lambda_j^*)\lambda_j^*\right)\theta_j\sigma^\eta + O(\eta^{1/3}),
\end{aligned}$$

where the last equality holds because $(\sigma^\eta)^2/\eta = \eta^{1/3}$. Similar inequality holds for the servers. So, by Eq (C.28), we have

$$\begin{aligned}
L^\eta =& \sum_{j\in M}\left(F_j'(\lambda_j^*)\lambda_j^* + F_j(\lambda_j^*)\right)\theta_j\sigma^\eta \Pr[q_j^{(c)} > \tau_{\max}^\eta] \\
&- \sum_{i\in N}\left(G_i'(\mu_i^*)\mu_i^* + G_i(\mu_i^*)\right)\phi_i\sigma^\eta \Pr[q_i^{(s)} > \tau_{\max}^\eta] + O(\eta^{1/3}) + \langle \mathbf{s}, \mathbb{E}[\mathbf{q}]\rangle.
\end{aligned}$$

By Lemma 2, we can simplify the first two terms in the above equation and get

$$L^\eta \le |E|^2 \max_{i\in N, j\in M}\{\theta_j, \phi_i\}\frac{(\sigma^\eta)^2}{\eta} + O(\eta^{1/3}) + \langle \mathbf{s}, \mathbb{E}[\mathbf{q}]\rangle. \tag{C.29}$$

Now, using Lemma 1, we can upper bound the expected queue length as

$$\begin{aligned}
\mathbb{E}\left[\left\langle \boldsymbol{\theta}, \mathbf{q}^{(c)}\right\rangle\right] + \mathbb{E}\left[\left\langle \boldsymbol{\phi}, \mathbf{q}^{(s)}\right\rangle\right] \le \tau_{\max}^\eta &\left(\sum_{j=1}^m \theta_j \Pr[q_j^{(c)} > \tau_{\max}^\eta] + \sum_{i=1}^n \phi_i \Pr[q_i^{(s)} > \tau_{\max}^\eta]\right) \\
&+ \frac{\eta}{\sigma^\eta}\left(\langle \mathbf{1}_n, \boldsymbol{\mu}^*\rangle + \langle \mathbf{1}_m, \boldsymbol{\lambda}^*\rangle\right),
\end{aligned}$$

which implies that

$$\min_{i\in N, j\in M}\left\{\frac{\phi_i}{s_i^{(s)}}, \frac{\theta_j}{s_j^{(c)}}\right\}\mathbb{E}[\langle \mathbf{s}, \mathbf{q}\rangle] \leq \tau^{\eta}_{\max}\left(\sum_{j=1}^{m}\theta_j + \sum_{i=1}^{n}\phi_i\right) + \frac{\eta}{\sigma^{\eta}}\left(\langle \mathbf{1}_n, \boldsymbol{\mu}^*\rangle + \langle \mathbf{1}_m, \boldsymbol{\lambda}^*\rangle\right).$$

By substituting $\tau^{\eta}_{\max} \leq \eta^{1/3}$, $\sigma^{\eta} = \eta^{2/3}$ and noting that $\min_{i\in N, j\in M}\{\phi_i/s_i^{(s)}, \theta_j/s_j^{(c)}\} > 0$, we have $\mathbb{E}[\langle \mathbf{s}, \mathbf{q}\rangle]=O(\eta^{1/3})$. By substituting this term in (C.29) and substituting $\sigma^{\eta} = \eta^{1/3}$, we have the desired result. $\square$

C.4 Lower Bounds

C.4.1 Proof of Theorem 7

First, we will present a lemma that provides a lower bound on the expected value of the sum of the queue length $\mathbb{E}\left[\langle \mathbf{1}_{n+m}, \mathbf{q}\rangle\right]$.

Lemma 18. *If Condition 1(a) is satisfied, then*

$$\mathbb{E}[\langle \mathbf{1}_{n+m}, \mathbf{q}\rangle] \geq \frac{\eta^{1-\beta}\langle \mathbf{1}_m, \boldsymbol{\lambda}^*\rangle}{2\left(\langle \mathbf{1}_n, \mathbf{N}\rangle + \langle \mathbf{1}_m, \mathbf{M}\rangle\right)} - 0.5. \tag{C.30}$$

In addition, if Condition 1(b) and 1(c) are satisfied, then there exists $\varepsilon > 0$ *and* $\mathscr{M} > 0$ *such that*

$$\sum_{j=1}^{m}\mathbb{E}\left[f_j^2\left(\frac{\mathbf{q}}{\eta^{\alpha}}\right)\right] + \sum_{i=1}^{n}\mathbb{E}\left[g_i^2\left(\frac{\mathbf{q}}{\eta^{\alpha}}\right)\right] \geq \varepsilon \quad \forall \eta > \mathscr{M}. \tag{C.31}$$

Proof of Lemma 18. We define the "imbalance" process of the two-sided queueing system as

$$z(k) = \left\langle \mathbf{1}_m, \mathbf{q}^{(c)}(k)\right\rangle - \left\langle \mathbf{1}_n, \mathbf{q}^{(s)}(k)\right\rangle. \tag{C.32}$$

Now, we will define a new DTMC $\{\tilde{z}(k) : k \in \mathbb{Z}_+\}$ and couple it with the uniformized

DTMC $\{\mathbf{q}(k), z(k) : k \in \mathbb{Z}_+\}$ with uniformization constant c such that $|z(k)| \geq \tilde{z}(k)$ for all $k \in \mathbb{Z}_+$ for all $\eta \geq 1$ if $z(0) = \tilde{z}(0)$.

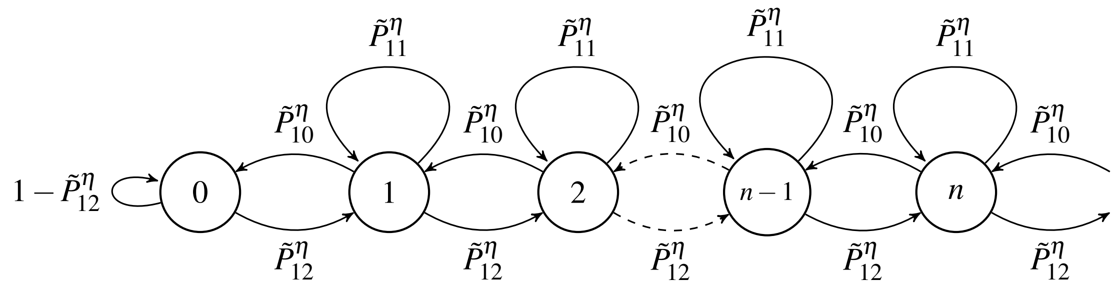


Figure C.3: Coupled Birth and Death Process

The state space of $\{\tilde{z}(k) : k \in \mathbb{Z}_+\}$ is $\mathbb{Z}_+$ and the transition matrix is given by

$$\tilde{P}^{\eta}_{ij} = \begin{cases} \tilde{P}^{\eta}_{10} \triangleq \left(\langle \mathbf{1}_n, \boldsymbol{\mu}^* \rangle + \left(\langle \mathbf{1}_n, \mathbf{N} \rangle + \langle \mathbf{1}_m, \mathbf{M} \rangle\right) \eta^{\beta-1}\right)/c & \text{if } j = i-1,\ \forall j \in \mathbb{Z}_+,\ \forall i > 0, \\ \tilde{P}^{\eta}_{12} \triangleq \left(\langle \mathbf{1}_m, \boldsymbol{\lambda}^* \rangle - \left(\langle \mathbf{1}_n, \mathbf{N} \rangle + \langle \mathbf{1}_m, \mathbf{M} \rangle\right) \eta^{\beta-1}\right)/c & \text{if } j = i+1,\ \forall j \in \mathbb{Z}_+,\ \forall i > 0, \\ \tilde{P}^{\eta}_{11} \triangleq 1 - \left(\langle \mathbf{1}_n, \boldsymbol{\mu}^* \rangle + \langle \mathbf{1}_m, \boldsymbol{\lambda}^* \rangle\right)/c & \text{if } j = i,\ \forall j \in \mathbb{Z}_+,\ \forall i > 0, \\ \left(\langle \mathbf{1}_m, \boldsymbol{\lambda}^* \rangle - \left(\langle \mathbf{1}_n, \mathbf{N} \rangle + \langle \mathbf{1}_m, \mathbf{M} \rangle\right) \eta^{\beta-1}\right)/c & \text{if } j = 1,\ i = 0, \\ 1 - \left(\langle \mathbf{1}_m, \boldsymbol{\lambda}^* \rangle - \left(\langle \mathbf{1}_n, \mathbf{N} \rangle + \langle \mathbf{1}_m, \mathbf{M} \rangle\right) \eta^{\beta-1}\right)/c & \text{if } j = 0,\ i = 0, \\ 0 & \text{otherwise}, \end{cases}$$

where c is the uniformization constant given by Definition 3. See Figure C.3 for an illustration of the new DTMC. Note that, under Condition 1(a), we have $\tilde{P}^{\eta}_{ij} \geq \Pr[|z(k+1)| = j,\ ||z(k)| = i,\ \mathbf{q}(k) = \bar{\mathbf{q}}]$ for all $\bar{\mathbf{q}} \in S$ if $j < i$ and $\tilde{P}^{\eta}_{ij} \leq \Pr[|z(k+1)| = j,\ ||z(k)| = i,\ \mathbf{q}(k) = \bar{\mathbf{q}}]$ for all $\bar{\mathbf{q}} \in S$ if $j > i$ for all $\eta > 0$. Thus, we can couple these system using a common source of randomness, such that $\tilde{z}(k) \leq |z(k)|$ for all $k \in \mathbb{Z}_+$ on each sample path. Hence, we have $\Pr[\tilde{z}(k) \leq \mathscr{K}] \geq \Pr[z(k) \leq \mathscr{K}]$ for all $k \geq 1$ and $\mathscr{K} \in \mathbb{R}$ Thus, in the limit as $k \to \infty$, we have $\Pr[\tilde{z}(\infty) \leq \mathscr{K}] \geq \Pr[z(\infty) \leq \mathscr{K}]$ where, $\tilde{z}(\infty)$ and $z(\infty)$ denotes random variables following the stationary distributions of $\{\tilde{z}(k) : k \geq 1\}$ and $\{z(k) : k \geq 1\}$, respectively. So,

we have

$$\mathbb{E}\left[\tilde{z}(\infty)\right] \leq \mathbb{E}\left[|z(\infty)|\right]. \tag{C.33}$$

The stationary distribution of $\{\tilde{z}(k) : k \geq 1\}$ is given by

$$\pi_i = \pi_0 \left(\frac{\tilde{P}_{12}^{\eta}}{\tilde{P}_{10}^{\eta}}\right)^i \quad \forall i > 0, \qquad \sum_{i=-\infty}^{\infty} \pi_i = 1.$$

Solving for π_i, we get

$$\pi_i = \left(1 - \frac{\tilde{P}_{12}^{\eta}}{\tilde{P}_{10}^{\eta}}\right)\left(\frac{\tilde{P}_{12}^{\eta}}{\tilde{P}_{10}^{\eta}}\right)^i \quad \forall i \in \mathbb{Z}_+.$$

Thus, we have

$$\mathbb{E}[\tilde{z}(\infty)] = \frac{\tilde{P}_{12}^{\eta}}{\tilde{P}_{10}^{\eta} - \tilde{P}_{12}^{\eta}} = \frac{\eta^{1-\beta}\left\langle \mathbf{1}_m, \boldsymbol{\lambda}^* \right\rangle}{2\left(\left\langle \mathbf{1}_n, \mathbf{N}\right\rangle + \left\langle \mathbf{1}_m, \mathbf{M}\right\rangle\right)} - 0.5. \tag{C.34}$$

where the second equality uses the definition of $\tilde{P}_{ij}^{\eta}$ above and the fact that $\langle \mathbf{1}_n, \boldsymbol{\mu}^* \rangle = \langle \mathbf{1}_m, \boldsymbol{\lambda}^* \rangle$. Finally, note that $|z| \leq \langle \mathbf{1}_{n+m}, \mathbf{q} \rangle$. Thus, we have

$$\mathbb{E}[\langle \mathbf{1}_{n+m}, \mathbf{q} \rangle] \ \geq\ \mathbb{E}[|z(\infty)|] \geq \mathbb{E}[\tilde{z}(\infty)] = \frac{\eta^{1-\beta}\left\langle \mathbf{1}_m, \boldsymbol{\lambda}^* \right\rangle}{2\left(\left\langle \mathbf{1}_n, \mathbf{N}\right\rangle + \left\langle \mathbf{1}_m, \mathbf{M}\right\rangle\right)} - 0.5,$$

where the second inequality uses (C.33) and the last equality uses (C.34). This completes the first part of the proof.

Next, as $\tilde{z}(\infty) \leq |z(\infty)|$ almost surely, for any constant $\mathscr{K}$, we have

$$\begin{aligned}\Pr[|z(\infty)| > \mathscr{K}] \geq \Pr[\tilde{z}(\infty) > \mathscr{K}] = \sum_{i=\mathscr{K}+1}^{\infty} \pi_i = \left(\frac{\tilde{P}_{12}^{\eta}}{\tilde{P}_{10}^{\eta}}\right)^{\mathscr{K}+1} \\ = \left(\frac{\langle \mathbf{1}_m, \boldsymbol{\lambda}^* \rangle - \left(\langle \mathbf{1}_n, \mathbf{N} \rangle + \langle \mathbf{1}_m, \mathbf{M} \rangle\right)\eta^{\beta-1}}{\langle \mathbf{1}_n, \boldsymbol{\mu}^* \rangle + \left(\langle \mathbf{1}_n, \mathbf{N} \rangle + \langle \mathbf{1}_m, \mathbf{M} \rangle\right)\eta^{\beta-1}}\right)^{\mathscr{K}+1}.\end{aligned}$$

Let $\mathscr{K} = (n+m)K\eta^{\alpha}$. When $\alpha = 0$, we have

$$\lim_{\eta\to\infty}\left(\frac{\langle \mathbf{1}_m, \boldsymbol{\lambda}^*\rangle - (\langle \mathbf{1}_n, \mathbf{N}\rangle + \langle \mathbf{1}_m, \mathbf{M}\rangle)\,\eta^{\beta-1}}{\langle \mathbf{1}_n, \boldsymbol{\mu}^*\rangle + (\langle \mathbf{1}_n, \mathbf{N}\rangle + \langle \mathbf{1}_m, \mathbf{M}\rangle)\,\eta^{\beta-1}}\right)^{(n+m)K+1} = 1^{(n+m)K+1} = 1.$$

For $\alpha > 0$, we have

$$\begin{aligned}
&\lim_{\eta\to\infty}\left(\frac{\langle \mathbf{1}_m, \boldsymbol{\lambda}^*\rangle - (\langle \mathbf{1}_n, \mathbf{N}\rangle + \langle \mathbf{1}_m, \mathbf{M}\rangle)\,\eta^{\beta-1}}{\langle \mathbf{1}_n, \boldsymbol{\mu}^*\rangle + (\langle \mathbf{1}_n, \mathbf{N}\rangle + \langle \mathbf{1}_m, \mathbf{M}\rangle)\,\eta^{\beta-1}}\right)^{(n+m)K\eta^{\alpha}} \\
&= \lim_{\eta\to\infty}\left[\left(1 - \frac{2\,(\langle \mathbf{1}_n, \mathbf{N}\rangle + \langle \mathbf{1}_m, \mathbf{M}\rangle)}{\langle \mathbf{1}_m, \boldsymbol{\mu}^*\rangle\,\eta^{1-\beta} + (\langle \mathbf{1}_n, \mathbf{N}\rangle + \langle \mathbf{1}_m, \mathbf{M}\rangle)}\right)^{\frac{\langle \mathbf{1}_n, \boldsymbol{\mu}^*\rangle\eta^{1-\beta}}{3(\langle \mathbf{1}_n, \mathbf{N}\rangle + \langle \mathbf{1}_m, \mathbf{M}\rangle)}}\right]^{K_0} \\
&\geq \lim_{\eta\to\infty}\left(1 - \frac{2}{3 + \frac{3(\langle \mathbf{1}_n, \mathbf{N}\rangle + \langle \mathbf{1}_m, \mathbf{M}\rangle)}{\langle \mathbf{1}_n, \boldsymbol{\mu}^*\rangle}\eta^{\beta-1}}\right)^{K_0} \\
&= \begin{cases} (1/3)^0 = 1 & \mathit{if}\ \alpha+\beta < 1, \\ b \triangleq \left(\frac{1}{3}\right)^{\frac{3K(n+m)(\langle \mathbf{1}_n, \mathbf{N}\rangle + \langle \mathbf{1}_m, \mathbf{M}\rangle)}{\langle \mathbf{1}_n, \boldsymbol{\mu}^*\rangle}} & \mathit{if}\ \alpha+\beta = 1, \end{cases}
\end{aligned}$$

where

$$K_0 = \frac{3K(n+m)\,(\langle \mathbf{1}_n, \mathbf{N}\rangle + \langle \mathbf{1}_m, \mathbf{M}\rangle)}{\langle \mathbf{1}_n, \boldsymbol{\mu}^*\rangle}\eta^{\alpha+\beta-1}$$

and the first inequality follows from the Bernoulli's inequality, which says that $(1+x)^r \geq$

$1+rx$ if $x \geq -2$. Thus, we have

$$
\begin{aligned}
&\lim_{\eta\to\infty}\left(\frac{\langle \mathbf{1}_m, \boldsymbol{\lambda}^*\rangle - \left(\langle \mathbf{1}_n, \mathbf{N}\rangle + \langle \mathbf{1}_m, \mathbf{M}\rangle\right)\eta^{\beta-1}}{\langle \mathbf{1}_n, \boldsymbol{\mu}^*\rangle + \left(\langle \mathbf{1}_n, \mathbf{N}\rangle + \langle \mathbf{1}_m, \mathbf{M}\rangle\right)\eta^{\beta-1}}\right)^{(n+m)K\eta^{\alpha}+1} \\
&= \lim_{\eta\to\infty}\left(\frac{\langle \mathbf{1}_m, \boldsymbol{\lambda}^*\rangle - \left(\langle \mathbf{1}_n, \mathbf{N}\rangle + \langle \mathbf{1}_m, \mathbf{M}\rangle\right)\eta^{\beta-1}}{\langle \mathbf{1}_n, \boldsymbol{\mu}^*\rangle + \left(\langle \mathbf{1}_n, \mathbf{N}\rangle + \langle \mathbf{1}_m, \mathbf{M}\rangle\right)\eta^{\beta-1}}\right)^{(n+m)K\eta^{\alpha}} \\
&\quad \cdot \lim_{\eta\to\infty}\left(\frac{\langle \mathbf{1}_m, \boldsymbol{\lambda}^*\rangle - \left(\langle \mathbf{1}_n, \mathbf{N}\rangle + \langle \mathbf{1}_m, \mathbf{M}\rangle\right)\eta^{\beta-1}}{\langle \mathbf{1}_n, \boldsymbol{\mu}^*\rangle + \left(\langle \mathbf{1}_n, \mathbf{N}\rangle + \langle \mathbf{1}_m, \mathbf{M}\rangle\right)\eta^{\beta-1}}\right) \\
&= \begin{cases} 1 & \text{if } \alpha+\beta < 1, \\ b & \text{if } \alpha+\beta = 1, \end{cases}
\end{aligned}
$$

where the first equality holds because the limit of the product of two sequences is equal to the product of the limits of the two sequences. Thus, when Condition 1(b) is satisfied, namely $\alpha+\beta \leq 1$, for any given $b\delta^2 > \varepsilon > 0$, there exists a constant $\mathscr{M}$ such that for all $\eta > \mathscr{M}$, it holds that

$$\Pr\left[|z(\infty)| > (n+m)K\eta^{\alpha}\right] \geq \frac{\varepsilon}{\delta^2}.$$

By the definition of z in Eq (C.34), the event

$$\left\{|z| > (n+m)K\eta^{\alpha}\right\} \subset \left\{\|\mathbf{q}\|_{\infty} > K\eta^{\alpha}\right\}.$$

Also, by Condition 1 (b), we have

$$\sum_{j=1}^{m} f_j^2\left(\frac{\mathbf{q}}{\eta^{\alpha}}\right) + \sum_{i=1}^{n} g_i^2\left(\frac{\mathbf{q}}{\eta^{\alpha}}\right) \geq \delta^2 \quad \forall \mathbf{q} : \|\mathbf{q}\|_{\infty} > k\eta^{\alpha}.$$

Thus for all $\eta > \mathscr{M}$, we have

$$\sum_{j=1}^{m} \mathbb{E}\left[f_j^2\left(\frac{\mathbf{q}}{\eta^{\alpha}}\right)\right] + \sum_{i=1}^{n} \mathbb{E}\left[g_i^2\left(\frac{\mathbf{q}}{\eta^{\alpha}}\right)\right] \geq \delta^2 \Pr\left[|z(\infty)| > (n+m)K\eta^{\alpha}\right] \geq \delta^2\frac{\varepsilon}{\delta^2} = \varepsilon.$$

$\square$

Claim 2. *It holds that*

$$A_j \triangleq -F_j'(\lambda_j^*) - \frac{1}{2}\lambda_j^* F_j''(\lambda_j^*) > 0, \quad \forall j \in M,$$
$$B_i \triangleq \; G_i'(\mu_i^*) + \frac{1}{2}\mu_i^* G_i''(\mu_i^*) > 0 \quad \forall i \in N.$$

There exists a constant $\eta' > 0$ *such that for all* $\eta > \eta'$ *and* $\tilde{c} \triangleq \min_{j\in M, i\in N}\left\{\frac{A_j}{\lambda_{j^*}}, \frac{B_i}{\mu_{i^*}}\right\}$*, it holds that*

$$\sup_{\mathbf{q}\in S} |F_j''(\tilde{\lambda}_j(\mathbf{q})) - F_j''(\lambda_j^*)| < \tilde{c} \quad \forall j \in M,$$
$$\sup_{\mathbf{q}\in S} |G_i''(\tilde{\mu}_i(\mathbf{q})) - G_i''(\mu_i^*)| < \tilde{c} \quad \forall i \in N.$$

Proof of Claim 2. As $-F_j(\lambda_j)\lambda_j$ for all $j \in M$ and $G_i(\mu_i)\mu_i$ for all $i \in N$ are strictly convex by Assumption 2, we have

$$\frac{d^2}{d\lambda_j^2}\left(-F_j(\lambda_j)\lambda_j\right) = 2A_j > 0 \;\; \forall j \in M, \quad \frac{d^2}{d\mu_i^2}\left(G_i(\mu_i)\mu_i\right) = 2B_i > 0 \;\; \forall i \in N.$$

Since $F_j''(.)$ is continuous by assumption, given $\tilde{c} = \min_{i,j}\{\frac{A_j}{\lambda_j^*}, \frac{B_i}{\mu_i^*}\} > 0$, there exists $\delta' > 0$ such that for any $l \in [\lambda_j^* - \delta', \lambda_j^* + \delta']$ we have

$$|F_j''(l) - F_j''(\lambda_j^*)| < \tilde{c}.$$

As $\beta < 1$, we have $\eta^{\beta-1} \to 0$ as $\eta \to \infty$. Consider η' such that for all $\eta > \eta'$ we have $|\eta^{\beta-1}| < \delta'/\max_{j\in M, i\in N}\{M_j, N_i\}$, which implies that $\tilde{\lambda}_j(\mathbf{q}) \in [\lambda_j^* - \delta', \lambda_j^* + \delta']$ for all $\mathbf{q}$. Thus,

$$\sup_{\mathbf{q}\in S} |F_j''(\tilde{\lambda}_j(\mathbf{q})) - F_j''(\lambda_j^*)| < \tilde{c}.$$

The proof for $G_i''(\mu)$ is similar. □

Proof of Theorem 7. Without loss of generality, we can assume that

$$\mathbb{E}[\langle \mathbf{1}_{n+m}, \mathbf{q} \rangle] < \infty; \tag{C.35}$$

otherwise, the holding cost will be infinity and the lower bound holds trivially. We calculate the profit loss and use Lemma 18 to lower bound the expected queue length. We have

$$\begin{aligned}
L^\eta = \eta \langle F(\boldsymbol{\lambda}^*), \boldsymbol{\lambda}^* \rangle - \eta \langle G(\boldsymbol{\mu}^*), \boldsymbol{\mu}^* \rangle - \mathbb{E}[\langle F^\eta(\boldsymbol{\lambda}^\eta), \boldsymbol{\lambda}^\eta \rangle - \langle G^\eta(\boldsymbol{\mu}^\eta), \boldsymbol{\mu}^\eta \rangle - s \langle \mathbf{1}_{n+m}, \mathbf{q} \rangle] \\
= & \eta \langle F(\boldsymbol{\lambda}^*), \boldsymbol{\lambda}^* \rangle - \eta \langle G(\boldsymbol{\mu}^*), \boldsymbol{\mu}^* \rangle \\
& - \mathbb{E}\left[\sum_{j=1}^{m} \left(\lambda_j^* + f_j\left(\frac{\mathbf{q}}{\eta^\alpha} \eta^{\beta-1} \right) \right) F_j \left(\lambda_j^* + f_j \left(\frac{\mathbf{q}}{\eta^\alpha} \eta^{\beta-1} \right) \right) \right. \\
& \left. - \sum_{i=1}^{n} \left(\mu_i^* + g_i \left(\frac{\mathbf{q}}{\eta^\alpha} \right) \eta^{\beta-1} \right) G_i \left(\mu_i^* + g_i \left(\frac{\mathbf{q}}{\eta^\alpha} \right) \eta^{\beta-1} \right) - \langle \mathbf{s}, \mathbf{q} \rangle \right],
\end{aligned} \tag{C.36}$$

where the second equality follows from the definition of the asymptotic regime (Definition 4) and the assumptions on the pricing policy Equation (4.10). Now, we will use Taylor's theorem to expand each function F_j and G_i individually. For the term involving F_j, we have

$$\begin{aligned}
& \mathbb{E}\left[\left(\eta \lambda_j^* + f_j \left(\frac{\mathbf{q}}{\eta^\alpha} \right) \eta^\beta \right) F_j \left(\lambda_j^* + f_j \left(\frac{\mathbf{q}}{\eta^\alpha} \right) \frac{\eta^\beta}{\eta} \right) \right] \\
& = \mathbb{E}\left[\left(\eta \lambda_j^* + f_j \left(\frac{\mathbf{q}}{\eta^\alpha} \right) \eta^\beta \right) \right. \\
& \quad \left. \cdot \left(F_j(\lambda_j^*) + f_j \left(\frac{\mathbf{q}}{\eta^\alpha} \right) \eta^{\beta-1} F_j'(\lambda_j^*) + f_j^2 \left(\frac{\mathbf{q}}{\eta^\alpha} \right) \frac{\eta^{2\beta-2}}{2} F_j'' \left(\tilde{\lambda}_j(\mathbf{q}) \right) \right) \right] \\
& = \eta \lambda_j^* F_j(\lambda_j^*) + \left(F_j'(\lambda_j^*) \lambda_j^* + F_j(\lambda_j^*) \right) \mathbb{E}\left[f_j \left(\frac{\mathbf{q}}{\eta^\alpha} \right) \right] \eta^\beta \\
& \quad + F_j'(\lambda_j^*) \mathbb{E}\left[f_j^2 \left(\frac{\mathbf{q}}{\eta^\alpha} \right) \right] \eta^{2\beta-1} + \lambda_j^* \mathbb{E}\left[f_j^2 \left(\frac{\mathbf{q}}{\eta^\alpha} \right) F_j'' \left(\tilde{\lambda}_j(\mathbf{q}) \right) \right] \frac{\eta^{2\beta-1}}{2} \\
& \quad + \mathbb{E}\left[f_j^3 \left(\frac{\mathbf{q}}{\eta^\alpha} \right) F_j'' \left(\tilde{\lambda}_j(\mathbf{q}) \right) \right] \frac{\eta^{3\beta-2}}{2},
\end{aligned}$$

where the first term follows from Taylor's theorem and $\tilde{\lambda}_j(\mathbf{q}) \in [\lambda_j^*, \lambda_j^* + f_j\left(\frac{\mathbf{q}}{\eta^\alpha}\right)\eta^{\beta-1}]$. Similarly, we can use the Taylor's theorem for all G_i in (C.36) by noting that $\tilde{\mu}_i(\mathbf{q}) \in [\mu_i^*, \mu_i^* + g_i\left(\frac{\mathbf{q}}{\eta^\alpha}\right)\eta^{\beta-1}]$ for all $i \in N$. Therefore, we have

$$
\begin{aligned}
L^\eta = & \underbrace{\sum_{i=1}^{n} \mathbb{E}\left[g_i\left(\frac{\mathbf{q}}{\eta^\alpha}\right)\eta^\beta\right]\left(\mu_i^* G_i'(\mu_i^*) + G_i(\mu_i^*)\right)}_{\mathscr{A}_1^{(s)}} \\
& \underbrace{-\sum_{j=1}^{m} \mathbb{E}\left[f_j\left(\frac{\mathbf{q}}{\eta^\alpha}\right)\eta^\beta\right]\left(\lambda_j^* F_j'(\lambda_j^*) + F_j(\lambda_j^*)\right)}_{\mathscr{A}_1^{(c)}} \\
& \underbrace{-\eta^{2\beta-1}\left(\sum_{j=1}^{m} F_j'(\lambda_j^*)\mathbb{E}\left[f_j^2\left(\frac{\mathbf{q}}{\eta^\alpha}\right)\right] + \sum_{j=1}^{m}\frac{\lambda_j^*}{2}\mathbb{E}\left[F_j''(\tilde{\lambda}_j(\mathbf{q}))f_j^2\left(\frac{\mathbf{q}}{\eta^\alpha}\right)\right]\right)}_{\mathscr{A}_2^{(c)}} \\
& \underbrace{+\eta^{2\beta-1}\left(\sum_{i=1}^{n} G_i'(\mu_i^*)\mathbb{E}\left[g_i^2\left(\frac{\mathbf{q}}{\eta^\alpha}\right)\right] + \sum_{i=1}^{n}\frac{\mu_i^*}{2}\mathbb{E}\left[G_i''(\tilde{\mu}(\mathbf{q}))g_i^2\left(\frac{\mathbf{q}}{\eta^\alpha}\right)\right]\right)}_{\mathscr{A}_2^{(s)}} \\
& \underbrace{-\frac{\eta^{3\beta-2}}{2}\left(\sum_{j=1}^{m}\mathbb{E}\left[F_j''(\tilde{\lambda}_j(\mathbf{q}))f_j^3\left(\frac{\mathbf{q}}{\eta^\alpha}\right)\right] - \sum_{i=1}^{n}\mathbb{E}\left[G_i''(\tilde{\mu}_i(\mathbf{q}))g_i^3\left(\frac{\mathbf{q}}{\eta^\alpha}\right)\right]\right)}_{\mathscr{A}_3} \\
& \underbrace{+\mathbb{E}[\langle \mathbf{s}, \mathbf{q}\rangle]}_{\mathscr{A}_4} \qquad \text{(C.37)}
\end{aligned}
$$

We now bound each of the terms in the above equation individually. Firstly, we have $\mathscr{A}_1^{(c)} + \mathscr{A}_1^{(s)} \geq 0$ by Claim 1 and (C.35). Next, the terms $\mathscr{A}_2^{(s)}$ and $\mathscr{A}_2^{(c)}$ can be simplified using Claim 2: for any $\eta > \eta'$, we have

$$
\begin{aligned}
& \mathscr{A}_2^{(c)} + \mathscr{A}_2^{(s)} \\
& \quad \geq \eta^{2\beta-1}\left(\sum_{j=1}^{m}\left(A_j - \frac{\lambda_j^*\tilde{c}}{2}\right)\mathbb{E}\left[f_j^2\left(\frac{\mathbf{q}}{\eta^\alpha}\right)\right] + \sum_{i=1}^{n}\left(B_i - \frac{\mu_i^*\tilde{c}}{2}\right)\mathbb{E}\left[g_i^2\left(\frac{\mathbf{q}}{\eta^\alpha}\right)\right]\right) \\
& \quad \geq \eta^{2\beta-1}\left(\sum_{j=1}^{m}\frac{A_j}{2}\mathbb{E}\left[f_j^2\left(\frac{\mathbf{q}}{\eta^\alpha}\right)\right] + \sum_{i=1}^{n}\frac{B_i}{2}\mathbb{E}\left[g_i^2\left(\frac{\mathbf{q}}{\eta^\alpha}\right)\right]\right),
\end{aligned}
$$

where the second inequality uses the definition of $\tilde{c}$. For the term $\mathscr{A}_3$, we have

$$\begin{aligned}\mathscr{A}_3 &= -\frac{\eta^{3\beta-2}}{2}\left(\sum_{j=1}^{m}\mathbb{E}\left[F_j''(\tilde{\lambda}_j(\mathbf{q}))f_j^3\left(\frac{\mathbf{q}}{\eta^{\alpha}}\right)\right]-\sum_{i=1}^{n}\mathbb{E}\left[G_i''(\tilde{\mu}_i(\mathbf{q}))g_i^3\left(\frac{\mathbf{q}}{\eta^{\alpha}}\right)\right]\right)\\ &\geq -\eta^{3\beta-2}\frac{\langle\mathbf{1}_m,|F''(\boldsymbol{\lambda}^*)|\rangle+\langle\mathbf{1}_n,|G''(\boldsymbol{\mu}^*)|\rangle+2\tilde{c}}{2}\left(\mathbb{E}\left[\left|f^3(\frac{\mathbf{q}}{\eta^{\alpha}})\right|\right]+\mathbb{E}\left[\left|g^3(\frac{\mathbf{q}}{\eta^{\alpha}})\right|\right]\right).\end{aligned}$$

Dividing both sides by $\eta^{2\beta-1}$, as $\beta<1$, we have

$$\begin{aligned}\frac{\mathscr{A}_3}{\eta^{2\beta-1}} &\geq -\eta^{\beta-1}\frac{\langle\mathbf{1}_m,|F''(\boldsymbol{\lambda}^*)|\rangle+\langle\mathbf{1}_n,|G''(\boldsymbol{\mu}^*)|\rangle+2\tilde{c}}{2}\\ &\qquad\cdot\left(\sum_{j=1}^{m}\mathbb{E}\left[\left|f_j^3(\frac{\mathbf{q}}{\eta^{\alpha}})\right|\right]+\sum_{i=1}^{n}\mathbb{E}\left[\left|g_i^3(\frac{\mathbf{q}}{\eta^{\alpha}})\right|\right]\right)\\ &\geq -\frac{\langle\mathbf{1}_m,|F''(\boldsymbol{\lambda}^*)|\rangle+\langle\mathbf{1}_n,|G''(\boldsymbol{\mu}^*)|\rangle+2\tilde{c}}{2}\\ &\qquad\cdot\left(\sum_{j=1}^{m}\mathbb{E}\left[\left|f_j^3(\frac{\mathbf{q}}{\eta^{\alpha}})\right|\right]+\sum_{i=1}^{n}\mathbb{E}\left[\left|g_i^3(\frac{\mathbf{q}}{\eta^{\alpha}})\right|\right]\right).\end{aligned}$$

For simplicity, we let $K_1=(\langle\mathbf{1}_m,|F''(\boldsymbol{\lambda}^*)|\rangle+\langle\mathbf{1}_n,|G''(\boldsymbol{\mu}^*)|\rangle+2\tilde{c})/2$. Finally, $\mathscr{A}_4$ can be lower bounded using Lemma 18. Combining everything, for all $\eta>\eta'$ (recall the definition of η' in Claim 2), we have

$$\begin{aligned}L^{\eta} &\geq \eta^{2\beta-1}\left(\sum_{j=1}^{m}\frac{A_j}{2}\mathbb{E}\left[f_j^2\left(\frac{\mathbf{q}}{\eta^{\alpha}}\right)\right]+\sum_{i=1}^{n}\frac{B_i}{2}\mathbb{E}\left[g_i^2\left(\frac{\mathbf{q}}{\eta^{\alpha}}\right)\right]\right)\\ &\quad-\eta^{2\beta-1}K_1\left(\sum_{j=1}^{m}\mathbb{E}\left[\left|f_j^3(\frac{\mathbf{q}}{\eta^{\alpha}})\right|\right]+\sum_{i=1}^{n}\mathbb{E}\left[\left|g_i^3(\frac{\mathbf{q}}{\eta^{\alpha}})\right|\right]\right)+\mathbb{E}\left[\langle\mathbf{s},\mathbf{q}\rangle\right].\end{aligned}\tag{C.38}$$

Let $K_2 \triangleq \left(4M^3+4N^3\right)K_1/\min_{i,j}\{A_j,B_i\}$. By Lemma 18 and Condition 1(a), we get

$$
\begin{aligned}
L^\eta &\geq \eta^{2\beta-1}\left(\frac{\min_{i,j}\left\{A_j,B_i\right\}K_2}{2}-\frac{\min_{i,j}\left\{A_j,B_i\right\}K_2}{4}\right)\\
&\quad+\frac{\min_{i\in N,j\in M}\{s_i^{(s)},s_j^{(c)}\}\eta^{1-\beta}\left\langle \mathbf{1}_m,\boldsymbol{\lambda}^*\right\rangle}{2\left(\left\langle \mathbf{1}_n,\mathbf{N}\right\rangle+\left\langle \mathbf{1}_m,\mathbf{M}\right\rangle\right)}-0.5\max_{i\in N,j\in M}\{s_i^{(s)},s_j^{(c)}\}\\
&\geq \inf_{\beta<1}\left\{\eta^{2\beta-1}\frac{\min_{i,j}\left\{A_j,B_i\right\}K_2}{4}+\frac{\min_{i\in N,j\in M}\{s_i^{(s)},s_j^{(c)}\}\eta^{1-\beta}\left\langle \mathbf{1}_m,\boldsymbol{\lambda}^*\right\rangle}{2\left(\left\langle \mathbf{1}_n,\mathbf{N}\right\rangle+\left\langle \mathbf{1}_m,\mathbf{M}\right\rangle\right)}\right\}\\
&\quad-0.5\max_{i\in N,j\in M}\{s_i^{(s)},s_j^{(c)}\}\\
&=\eta^{1/3}\left(\frac{\min_{i,j}\left\{A_j,B_i\right\}K_2}{4}+\frac{\min_{i\in N,j\in M}\{s_i^{(s)},s_j^{(c)}\}\left\langle \mathbf{1}_m,\boldsymbol{\lambda}^*\right\rangle}{2\left(\left\langle \mathbf{1}_n,\mathbf{N}\right\rangle+\left\langle \mathbf{1}_m,\mathbf{M}\right\rangle\right)}\right)\\
&\quad-0.5\max_{i\in N,j\in M}\{s_i^{(s)},s_j^{(c)}\},
\end{aligned}
$$

where the last equality follows as the coefficient of the terms $\eta^{2\beta-1}$ and $\eta^{1-\beta}$ are strictly positive and the infimum is achieved when $2\beta-1=1-\beta$. Let

$$
K=\frac{\min_{i,j}\left\{A_j,B_i\right\}K_2}{4}+\frac{\min_{i\in N,j\in M}\{s_i^{(s)},s_j^{(c)}\}\left\langle \mathbf{1}_m,\boldsymbol{\lambda}^*\right\rangle}{2\left(\left\langle \mathbf{1}_n,\mathbf{N}\right\rangle+\left\langle \mathbf{1}_m,\mathbf{M}\right\rangle\right)},
$$

and the proof is complete. □

C.4.2 Proof of Proposition 7

Lemma 19. *Under the hypothesis of Proposition 7, if $\alpha+\beta>1$, there exists $\tilde{c}_1>0$, $\tilde{c}_2>0$ such that for all $\eta\geq 1$, it holds that*

$$
\begin{aligned}
&\mathbb{E}[q^{(s)}+q^{(c)}]\geq \max\left\{\frac{1}{2}\eta^\alpha,\frac{\eta^{1-\beta}\lambda^*}{2(\theta+\phi)}-0.5\right\}\\
&\mathbb{E}\left[f^2\left(\frac{\mathbf{q}}{\eta^\alpha}\right)+g^2\left(\frac{\mathbf{q}}{\eta^\alpha}\right)\right]\geq \tilde{c}_1\eta^{1-\alpha-\beta}\\
&\mathbb{E}\left[\left|f^3\left(\frac{\mathbf{q}}{\eta^\alpha}\right)\right|+\left|g^3\left(\frac{\mathbf{q}}{\eta^\alpha}\right)\right|\right]\leq \tilde{c}_2\eta^{1-\alpha-\beta},
\end{aligned}
$$

where $f(\mathbf{q}/\eta^{\alpha}) = -\theta \mathbb{1}_{\{q^{(c)} \geq \eta^{\alpha}\}}$ *and* $g(\mathbf{q}/\eta^{\alpha}) = -\phi \mathbb{1}_{\{q^{(s)} \geq \eta^{\alpha}\}}$.

Proof of Lemma 19. Recall that the two-price policy for a single-link two-sided queue has the form

$$\lambda^{\eta}(q) = \begin{cases} \eta\lambda^* & \text{if } q \leq \lceil \eta^{\alpha} \rceil \\ \eta\lambda^* - \theta\eta^{\beta} & \text{otherwise} \end{cases} \quad \mu^{\eta}(q) = \begin{cases} \eta\mu^* & \text{if } q \geq -\lceil \eta^{\alpha} \rceil \\ \eta\mu^* - \phi\eta^{\beta} & \text{otherwise,} \end{cases}$$

where $q = q^{(c)} - q^{(s)}$. Also, we have $\min\{q^{(c)}, q^{(s)}\} = 0$ for single-link queues. Consider a birth and death process denoted by $\{\tilde{q}(t) : t \geq 0\}$ with the state space $\mathbb{Z}$ and the transition rates given by

$$\tilde{\lambda}^{\eta}(q) = \begin{cases} \eta\lambda^* & \text{if } q \leq \lceil \eta^{\alpha} \rceil \\ 0 & \text{otherwise} \end{cases} \quad \tilde{\mu}^{\eta}(q) = \begin{cases} \eta\mu^* & \text{if } q \geq -\lceil \eta^{\alpha} \rceil \\ 0 & \text{otherwise.} \end{cases}$$

Because $\tilde{\lambda}(q) \leq \lambda(q)$ and $\tilde{\mu}(q) \geq \mu(q)$ for all $q \geq 0$ and $\tilde{\lambda}(q) \geq \lambda(q)$ and $\tilde{\mu}(q) \leq \mu(q)$ for all $q \geq 0$, we can couple the two systems such that $|\tilde{q}(t)| \leq |q(t)|$ for all $t \geq 1$. Thus, we have $\Pr[|q(t)| \leq \mathscr{K}] \leq \Pr[|\tilde{q}(t)| \leq \mathscr{K}]$ for all $t \in \mathbb{Z}_+$. Since $\{q(t) : t \geq 0\}$ is irreducible and positive recurrent by assumption and $\{\tilde{q}(t) : t \geq 0\}$ has only one irreducible and positive recurrent class by construction, the stationary distributions exist for both q and $\tilde{q}$. We have

$$\Pr[|q(\infty)| \leq \mathscr{K}] \leq \Pr[|\tilde{q}(\infty)| \leq \mathscr{K}] \; \forall \mathscr{K} \in \mathbb{R} \; \Rightarrow \mathbb{E}[|q(\infty)|] \geq \mathbb{E}[|\tilde{q}(\infty)|],$$

where $q(\infty)$ and $\tilde{q}(\infty)$ are random variables following the stationary distributions of $\{q(t) : t \geq 0\}$ and $\{\tilde{q}(t) : t \geq 0\}$, respectively. It is easy to verify that $\mathbb{E}[|\tilde{q}(\infty)|] = \lceil \eta^{\alpha} \rceil / 2$. Thus, we have

$$\mathbb{E}[|q^{(s)}(\infty) + q^{(c)}(\infty)] = \mathbb{E}[|q(\infty)|] \geq \mathbb{E}[|\tilde{q}(\infty)|] \geq \frac{\eta^{\alpha}}{2}.$$

Also, by Eq (4.12), the two-price policy satisfies Condition 1(a) and 1(c). By Lemma 18, we have

$$\mathbb{E}[|q(\infty)|] = \mathbb{E}[q^{(s)}(\infty) + q^{(c)}(\infty)] \geq \frac{\eta^{1-\beta}\lambda^*}{2(\theta+\phi)} - 0.5,$$

where $(q^{(s)}(\infty), q^{(c)}(\infty))$ are random variables which follows the stationary distributions of $\{q^{(s)}(t), q^{(c)}(t) : t \geq 0\}$. The first equality follows as $|q(t)| = q^{(s)}(t) + q^{(c)}(t)$ *a.s.* This proves the first inequality in the lemma.

Now, we will lower bound $\Pr[|q| \geq \eta^\alpha]$ for a given two-price policy. Since $\lambda^* = \mu^*$, the stationary distribution under a given two-price policy denoted by $\{\pi_i\}_{i\in\mathbb{Z}}$ is given by

$$\pi_k = \begin{cases} \pi_0 & \forall k \leq \lceil \eta^\alpha \rceil, k \geq -\lceil \eta^\alpha \rceil \\ \pi_0 \left(1 - \frac{\theta\eta^\beta}{\eta\mu^*}\right)^{k-\lceil \eta^\alpha \rceil} & \forall k > \lceil \eta^\alpha \rceil \\ \pi_0 \left(1 - \frac{\phi\eta^\beta}{\eta\lambda^*}\right)^{-k+\lceil \eta^\alpha \rceil} & \forall k < -\lceil \eta^\alpha \rceil. \end{cases}$$

This implies that

$$\begin{aligned} \Pr[|q(\infty)| \geq \eta^\alpha] &= \frac{\eta^{1-\beta}(\mu^*/\theta + \lambda^*/\phi)}{\eta^{1-\beta}(\mu^*/\theta + \lambda^*/\phi) + 2\lceil \eta^\alpha \rceil - 1} \\ &\geq \frac{(\mu^*/\theta + \lambda^*/\phi)}{(\mu^*/\theta + \lambda^*/\phi) + 3}\eta^{1-\alpha-\beta} \triangleq \frac{\tilde{c}_1}{\min\{\theta^2, \phi^2\}}\eta^{1-\alpha-\beta} \quad \forall \eta \geq 1, \end{aligned}$$

where the inequality follows as $1 - \alpha - \beta < 0$. This implies that

$$\begin{aligned} \mathbb{E}\left[f^2\left(\frac{\mathbf{q}}{\eta^\alpha}\right)\right] + \mathbb{E}\left[g^2\left(\frac{\mathbf{q}}{\eta^\alpha}\right)\right] &= \theta^2 \Pr\left[q^{(c)}(\infty) \geq \eta^\alpha\right] + \phi^2 \Pr\left[q^{(s)}(\infty) \geq \eta^\alpha\right] \\ &\geq \min\{\theta^2, \phi^2\} \Pr\left[|q(\infty)| \geq \eta^\alpha\right] \geq \tilde{c}_1 \eta^{1-\alpha-\beta}. \end{aligned}$$

Thus, the second inequality of the lemma is proved.

Similarly, we have

$$\begin{aligned}\Pr[|q(\infty)| \geq \eta^{\alpha}] \leq & \frac{\eta^{1-\alpha-\beta}(\mu^*/\theta + \lambda^*/\phi)}{\eta^{1-\alpha-\beta}(\mu^*/\theta + \lambda^*/\phi) + 2 - \eta^{-\alpha}} \\ & \leq (\mu^*/\theta + \lambda^*/\phi)\eta^{1-\alpha-\beta} \triangleq \frac{\tilde{c}_2}{\max\{\theta^3, \phi^3\}} \eta^{1-\alpha-\beta} \quad \forall \eta \geq 1,\end{aligned}$$

where we replace $\lceil \eta^{\alpha} \rceil$ by η^{α} in the first inequality, and the second inequality follows as $\alpha \geq 0$ and $1 - \alpha - \beta < 0$. This implies that

$$\begin{aligned}\mathbb{E}\left[\left|f^3\left(\frac{\mathbf{q}}{\eta^{\alpha}}\right)\right|\right] + \mathbb{E}\left[\left|g^3\left(\frac{\mathbf{q}}{\eta^{\alpha}}\right)\right|\right] &= \theta^3 \Pr\left[q^{(c)}(\infty) \geq \eta^{\alpha}\right] + \phi^3 \Pr\left[q^{(s)}(\infty) \geq \eta^{\alpha}\right] \\ &\leq \max\{\theta^3, \phi^3\} \Pr\left[|q(\infty)| \geq \eta^{\alpha}\right] \leq \tilde{c}_2 \eta^{1-\alpha-\beta}.\end{aligned}$$

This proves the last inequality of the lemma. $\square$

Proof of Proposition 7. When $\alpha + \beta \leq 1$, by Theorem 7, we know that $L^{\eta} \geq K\eta^{1/3}$. So we will only prove Proposition 7 when $\alpha + \beta > 1$. As the two-price policy is a special case of the general pricing policy given by Eqs (4.10, 4.11), it holds by Eq (C.38) that

$$\begin{aligned}L^{\eta} \geq & \eta^{2\beta-1}\left(\frac{A}{2}\mathbb{E}\left[f^2\left(\frac{\mathbf{q}}{\eta^{\alpha}}\right)\right] + \frac{B}{2}\mathbb{E}\left[g^2\left(\frac{\mathbf{q}}{\eta^{\alpha}}\right)\right]\right) \\ & - \eta^{2\beta-1}K_1\left(\mathbb{E}\left[\left|f^3(\frac{\mathbf{q}}{\eta^{\alpha}})\right|\right] + \mathbb{E}\left[\left|g^3(\frac{\mathbf{q}}{\eta^{\alpha}})\right|\right]\right) + \mathbb{E}\left[\langle \mathbf{s}, \mathbf{q} \rangle\right] \\ \geq & \eta^{\beta-\alpha}\left(\min\{A, B\}\frac{\tilde{c}_1}{2} - \tilde{c}_2 K_1\right) + \min_{i \in N, j \in M}\{s_i^{(s)}, s_j^{(c)}\} \max\left\{\frac{1}{2}\eta^{\alpha}, \frac{\eta^{1-\beta}\lambda^*}{2(\theta+\phi)} - 0.5\right\}.\end{aligned}$$

Letting $K_2 = \min\{A, B\}K_1\tilde{c}_1/(4\tilde{c}_2)$ and optimizing over all possible values of α and β, we get

$$L^\eta \geq \inf_{\alpha\geq 0,\ 1-\beta<\alpha} \left\{ \eta^{\beta-\alpha} \min\{A,B\} \frac{K_2}{4} + \min_{i\in N, j\in M} \{s_i^{(s)}, s_j^{(c)}\} \max\left\{ \frac{1}{2}\eta^\alpha, \frac{\eta^{1-\beta}\lambda^*}{2(\theta+\phi)} - 0.5 \right\} \right\}$$

$$= \left(\min\{A,B\}\frac{K_2}{4} + \min_{i\in N, j\in M}\{s_i^{(s)}, s_j^{(c)}\} \max\left\{ \frac{1}{2}, \frac{\lambda^*}{2(\theta+\phi)} - \frac{1}{2\eta^{1/3}} \right\} \right) \eta^{1/3}, \quad \text{(C.39)}$$

where the infimum is achieved when $\alpha = 1/3$, $\beta = 2/3$. □

C.5 Further Analysis on Max-Weight Matching

Proof of Lemma 3. By the CRP condition, there exists a constant $\delta > 0$ such that for all $J \subsetneq N$ and for all $I \subsetneq N$, it holds that

$$\sum_{j\in J} \lambda_j^* < \sum_{i:\exists j\in J, (i,j)\in E} \mu_i^* - \delta, \quad \sum_{i\in I} \mu_i^* < \sum_{j:\exists i\in I, (i,j)\in E} \lambda_j^* - \delta. \quad \text{(C.40)}$$

We claim that there exists $\chi \in \mathbb{R}_+^{m\times n}$ satisfying the following equalities:

$$\lambda_j^* - \frac{\delta}{n^2}|N(j)| = \sum_{i=1}^{n} \chi_{ij}, \quad \forall j \in M \quad \text{(C.41a)}$$

$$\mu_i^* - \frac{\delta}{n^2}|N(i)| = \sum_{j=1}^{m} \chi_{ij}, \quad \forall i \in N \quad \text{(C.41b)}$$

$$\chi_{ij} = 0, \quad \forall (i,j) \notin E. \quad \text{(C.41c)}$$

Such a χ exists because the left-hand sides $(\lambda_j^* - \frac{\delta}{n^2}|N(j)|)_{j=1}^m$ and $(\mu_i^* - \frac{\delta}{n^2}|N(i)|)_{i=1}^n$ satisfy the Hall's condition, which is implied by (C.40).

Let $\chi_{ij}^* = \frac{\delta}{n^2} + \chi_{ij}$ for all $(i,j) \in E$ and $\chi_{ij}^* = 0$ otherwise. Substituting χ for χ^* in

(C.41) gives

$$\lambda_j^* = \sum_{i=1}^{n} \chi_{ij}^*, \quad \forall j \in M \tag{C.42a}$$

$$\mu_i^* = \sum_{j=1}^{m} \chi_{ij}^*, \quad \forall i \in N \tag{C.42b}$$

$$\chi_{ij}^* = 0, \quad \forall (i,j) \notin E. \tag{C.42c}$$

Note that (C.42) is the same as the constraints in the fluid problem (4.6), so $(\boldsymbol{\lambda}^*, \boldsymbol{\mu}^*, \boldsymbol{\chi}^*)$ is a feasible solution. In addition, $\langle F(\boldsymbol{\lambda}^*), \boldsymbol{\lambda}^* \rangle - \langle G(\boldsymbol{\mu}^*), \boldsymbol{\mu}^* \rangle$ is the optimal objective function value, so $(\boldsymbol{\lambda}^*, \boldsymbol{\mu}^*, \boldsymbol{\chi}^*)$ is an optimal solution. As $\delta > 0$, we have $\chi_{ij}^* > 0$ for all $(i,j) \in E$. □

To prove the state space collapse, we consider the Lyapunov function

$$U(\mathbf{q}) = \sqrt{\langle \mathbf{1}_{2n}, \mathbf{q}^2 \rangle - \frac{z^2}{n}}. \tag{C.43}$$

(Recall the definition of z from Eq (C.32).) In addition, define the following vectors of size n: $\mathbf{q}_{\parallel}^{(s)} \triangleq \frac{1}{n} \left\langle \mathbf{1}_n, \mathbf{q}^{(s)} \right\rangle \mathbf{1}_n$, $\mathbf{q}_{\perp}^{(s)} \triangleq \mathbf{q}^{(s)} - \mathbf{q}_{\parallel}^{(s)}$ and $\mathbf{q}_{\parallel}^{(c)} \triangleq \frac{1}{n} \left\langle \mathbf{1}_n, \mathbf{q}^{(c)} \right\rangle \mathbf{1}_n$, $\mathbf{q}_{\perp}^{(c)} \triangleq \mathbf{q}^{(c)} - \mathbf{q}_{\parallel}^{(c)}$. We will first prove the following lemmas, which will assist us in proving Proposition 8 and Proposition 9.

Lemma 20. *Under any pricing policy and max-weight matching policy, for all* $\mathbf{q} \in S$*, we have*

$$U(\mathbf{q}) = \|\mathbf{q}_{\perp}^{(s)} - \mathbf{q}_{\perp}^{(c)}\|.$$

Proof of Lemma 20. With elementary algebra, we can show

$$\begin{aligned}
U^2(\mathbf{q}) &= \left\langle \mathbf{1}_n, (\mathbf{q}^{(s)})^2 \right\rangle - \frac{1}{n}\left\langle \mathbf{1}_n, \mathbf{q}^{(s)} \right\rangle^2 + \left\langle \mathbf{1}_n, (\mathbf{q}^{(c)})^2 \right\rangle - \frac{1}{n}\left\langle \mathbf{1}_n, \mathbf{q}^{(c)} \right\rangle^2 \\
&\quad + \frac{2}{n}\left\langle \mathbf{1}_n, \mathbf{q}^{(s)} \right\rangle \left\langle \mathbf{1}_n, \mathbf{q}^{(c)} \right\rangle \\
&= \left\langle \mathbf{1}_n, (\mathbf{q}_\perp^{(s)})^2 \right\rangle + \left\langle \mathbf{1}_n, (\mathbf{q}_\perp^{(c)})^2 \right\rangle + \frac{2}{n}\left\langle \mathbf{1}_n, \mathbf{q}^{(s)} \right\rangle \left\langle \mathbf{1}_n, \mathbf{q}^{(c)} \right\rangle \\
&= \left\langle \mathbf{1}_n, (\mathbf{q}_\perp^{(s)})^2 \right\rangle + \left\langle \mathbf{1}_n, (\mathbf{q}_\perp^{(c)})^2 \right\rangle - 2\left\langle \mathbf{q}_\perp^{(s)}, \mathbf{q}_\perp^{(c)} \right\rangle \\
&= \|\mathbf{q}_\perp^{(s)} - \mathbf{q}_\perp^{(c)}\|^2,
\end{aligned}$$

where the third equality above holds because

$$\begin{aligned}
\left\langle \mathbf{q}_\perp^{(s)}, \mathbf{q}_\perp^{(c)} \right\rangle &= \sum_{i=1}^n \left(q_i^{(s)} - \frac{1}{n}\left\langle \mathbf{1}_n, \mathbf{q}^{(s)} \right\rangle \right) \left(q_i^{(c)} - \frac{1}{n}\left\langle \mathbf{1}_n, \mathbf{q}^{(c)} \right\rangle \right) \\
&= \sum_{i=1}^n q_i^{(s)} q_i^{(c)} - \frac{1}{n}\left\langle \mathbf{1}_n, \mathbf{q}^{(s)} \right\rangle \left\langle \mathbf{1}_n, \mathbf{q}^{(c)} \right\rangle \\
&= -\frac{1}{n}\left\langle \mathbf{1}_n, \mathbf{q}^{(s)} \right\rangle \left\langle \mathbf{1}_n, \mathbf{q}^{(c)} \right\rangle.
\end{aligned}$$

The last equality above follows because E contains edges connecting type i server to type i customer by Condition 3. Thus, for the max-weight matching policy, either $q_i^{(s)} = 0$ or $q_i^{(c)} = 0$. $\square$

Lemma 21. *Under any pricing policy and max-weight matching policy, for all* $\mathbf{q} \in \mathbb{Z}_+^{m+n}$, *we have*

$$T(\mathbf{q}) := \sum_{(i,j)\in E} \left(q_i^{(s)} + q_j^{(c)} - \max_{i':(i',j)\in E} q_{i'}^{(s)} - \max_{j':(i,j')\in E} q_{j'}^{(c)} \right) \leq -\frac{1}{n}\sqrt{\langle \mathbf{1}_{2n}, \mathbf{q}^2 \rangle - \frac{z^2}{n}}.$$

Proof of Lemma 21. By the definition of $T(\mathbf{q})$, we have

$$T(\mathbf{q}) = \underbrace{\sum_{(i,j)\in E} \left(q_i^{(s)} - \max_{i':(i',j)\in E} q_{i'}^{(s)} \right)}_{T_1} + \underbrace{\sum_{(i,j)\in E} \left(q_j^{(c)} - \max_{j':(i,j')\in E} q_{j'}^{(c)} \right)}_{T_2}.$$

As the graph E is connected (implied by the CRP condition), consider a path P from $i_{\min} = \min_{i\in N} q_i^{(s)}$ to $i_{\max} = \max_{i\in N} q_i^{(s)}$. We have

$$\begin{aligned}
T_1 &= \sum_{(i,j)\in E} \left(q_i^{(s)} - \max_{i':(i',j)\in E} q_{i'}^{(s)} \right) \\
&= \sum_{(i,j)\in P} \left(q_i^{(s)} - \max_{i':(i',j)\in E} q_{i'}^{(s)} \right) + \sum_{(i,j)\in E\setminus P} \left(q_i^{(s)} - \max_{i':(i',j)\in E} q_{i'}^{(s)} \right) \\
&\le \sum_{(i,j)\in P} \left(q_i^{(s)} - \max_{i':(i',j)\in E} q_{i'}^{(s)} \right) \le q_{i_{\min}}^{(s)} - q_{i_{\max}}^{(s)},
\end{aligned}$$

where the last inequality follows from telescoping sum. Similarly, $T_2 \le q_{j_{\min}}^{(c)} - q_{j_{\max}}^{(c)}$. Thus, we have

$$T \le q_{i_{\min}}^{(s)} + q_{j_{\min}}^{(c)} - q_{i_{\max}}^{(s)} - q_{j_{\max}}^{(c)}.$$

Note that, $q_{i_{\min}}^{(s)}$ and $q_{j_{\min}}^{(c)}$ cannot both be non-zero by the definition of the max-weight algorithm. Without loss of generality, we assume $q_{j_{\min}}^{(c)} = 0$. We consider two cases:

Case I: $q_{i_{\min}}^{(s)} = 0$. We have

$$T(\mathbf{q}) \le -\left(q_{i_{\max}}^{(s)} + q_{j_{\max}}^{(c)} \right) \le -\frac{1}{n} \langle \mathbf{1}_{2n}, \mathbf{q} \rangle \le -\frac{1}{n} \|\mathbf{q}\| \le -\frac{1}{n} U(\mathbf{q}).$$

Case II: $q_{i_{\min}}^{(s)} > 0$. This implies that $q_i^{(s)} > 0$ for all $i \in N$, so we must have $q_j^{(c)} = 0$ for all $j \in N$ by the definition of the max-weight algorithm. Thus, we have

$$T(\mathbf{q}) \le q_{i_{\min}}^{(s)} - q_{i_{\max}}^{(s)} \le \frac{1}{n} \left\langle \mathbf{1}_n, \mathbf{q}^{(s)} \right\rangle - q_{i_{\max}}^{(s)} \le -\frac{1}{n} \|q_\perp^{(s)}\| = -\frac{1}{n} \|q_\perp^{(s)} - q_\perp^{(c)}\| = -\frac{1}{n} U(\mathbf{q}).$$

$\square$

To prepare for the next lemma, recall the definition

$$V(\mathbf{q}) = \left\langle \mathbf{1}_{2n}, \mathbf{q}^2 \right\rangle, \; V^{(c)}(\mathbf{q}) = \left\langle \mathbf{1}_n, (\mathbf{q}^{(c)})^2 \right\rangle, \; V^{(s)}(\mathbf{q}) = \left\langle \mathbf{1}_n, (\mathbf{q}^{(s)})^2 \right\rangle.$$

We also define

$$W_z(\mathbf{q}) = \frac{z^2}{n}. \tag{C.44}$$

Lemma 22. *Under any pricing policy and max-weight matching policy, the drift of $U(\mathbf{q})$, defined as $\Delta U(\mathbf{q}) = (U(\mathbf{q}(k+1)) - U(\mathbf{q}(k)))\mathbb{1}_{\{\mathbf{q}(k)=\mathbf{q}\}}$, satisfies*

$$|\Delta U(\mathbf{q})| \leq 4, \quad \Delta U(\mathbf{q}) \leq \frac{1}{2U(\mathbf{q})}(\Delta V(\mathbf{q}) - \Delta W_z(\mathbf{q})).$$

Proof of Lemma 22. By expanding the drift term, we have

$$\begin{aligned}
|\Delta U(\mathbf{q})| &= |U(\mathbf{q}(k+1)) - U(\mathbf{q}(k))|\mathbb{1}_{\mathbf{q}(k)=\mathbf{q}} \\
&\overset{(a)}{=} \left| \|\mathbf{q}_\perp^{(s)}(k+1) - \mathbf{q}_\perp^{(c)}(k+1)\| - \|\mathbf{q}_\perp^{(s)}(k) - \mathbf{q}_\perp^{(c)}(k)\| \right| \mathbb{1}_{\mathbf{q}(k)=\mathbf{q}} \\
&\overset{(b)}{\leq} \|\mathbf{q}_\perp^{(s)}(k+1) - \mathbf{q}_\perp^{(c)}(k+1) - \mathbf{q}_\perp^{(s)}(k) + \mathbf{q}_\perp^{(c)}(k)\|\mathbb{1}_{\mathbf{q}(k)=\mathbf{q}} \\
&= \|\mathbf{q}^{(s)}(k+1) - \mathbf{q}^{(s)}(k) - \mathbf{q}_\parallel^{(s)}(k+1) + \mathbf{q}_\parallel^{(s)}(k) - \mathbf{q}^{(c)}(k+1) + \mathbf{q}^{(c)}(k) \\
&\qquad + \mathbf{q}_\parallel^{(c)}(k+1) - \mathbf{q}_\parallel^{(c)}(k)\|\mathbb{1}_{\mathbf{q}(k)=\mathbf{q}} \\
&\overset{(c)}{\leq} \Big(\|\mathbf{q}^{(s)}(k+1) - \mathbf{q}^{(s)}(k)\| + \|\mathbf{q}_\parallel^{(s)}(k+1) - \mathbf{q}_\parallel^{(s)}(k)\| + \|\mathbf{q}^{(c)}(k+1) - \mathbf{q}^{(c)}(k)\| \\
&\qquad + \|\mathbf{q}_\parallel^{(c)}(k+1) - \mathbf{q}_\parallel^{(c)}(k)\| \Big) \mathbb{1}_{\mathbf{q}(k)=\mathbf{q}} \\
&\overset{(d)}{\leq} 4\|\mathbf{q}(k+1) - \mathbf{q}(k)\|\mathbb{1}_{\mathbf{q}(k)=\mathbf{q}} \\
&\overset{(e)}{\leq} 4,
\end{aligned} \tag{C.45}$$

where (a) follows from Lemma 20; (b) and (c) follow by the triangle inequality; (d) follows because projection onto a subspace is non-expansive and $\|\mathbf{q}^{(l)}(k+1) - \mathbf{q}^{(l)}(k)\| \leq \|\mathbf{q}(k+1) - \mathbf{q}(k)\|$ for $l \in \{1,2\}$; finally, (e) follows as there can be at most one arrival and one matching in one time epoch of the unformized DTMC. This proves the first part of the lemma.

Now we will show the second part of the lemma. Expanding the drift term, we have

$$
\begin{aligned}
\Delta U(\mathbf{q}) &= \left(\|\mathbf{q}_\perp^{(s)}(k+1) - \mathbf{q}_\perp^{(c)}(k+1)\| - \|\mathbf{q}_\perp^{(s)}(k) - \mathbf{q}_\perp^{(c)}(k)\|\right) \mathbb{1}_{\mathbf{q}(k)=\mathbf{q}} \\
&= \left(\sqrt{\|\mathbf{q}_\perp^{(s)}(k+1) - \mathbf{q}_\perp^{(c)}(k+1)\|^2} - \sqrt{\|\mathbf{q}_\perp^{(s)}(k) - \mathbf{q}_\perp^{(c)}(k)\|^2}\right) \mathbb{1}_{\mathbf{q}(k)=\mathbf{q}} \\
&\leq \frac{1}{2\|\mathbf{q}_\perp^{(s)}(k) - \mathbf{q}_\perp^{(c)}(k)\|} \\
&\quad \cdot \left(\|\mathbf{q}_\perp^{(s)}(k+1) - \mathbf{q}_\perp^{(c)}(k+1)\|^2 - \|\mathbf{q}_\perp^{(s)}(k) - \mathbf{q}_\perp^{(c)}(k)\|^2\right) \mathbb{1}_{\mathbf{q}(k)=\mathbf{q}} \\
&= \frac{1}{2\|\mathbf{q}_\perp^{(s)}(k) - \mathbf{q}_\perp^{(c)}(k)\|} \\
&\quad \cdot \left(\langle \mathbf{1}_{2n}, \mathbf{q}^2(k+1)\rangle - \langle \mathbf{1}_{2n}, \mathbf{q}^2(k)\rangle - \frac{z^2(k+1)}{n} + \frac{z^2(k)}{n}\right) \mathbb{1}_{\mathbf{q}(k)=\mathbf{q}} \\
&= \frac{1}{2\|\mathbf{q}_\perp^{(s)} - \mathbf{q}_\perp^{(c)}\|} \left(\Delta V(\mathbf{q}) - \Delta W_z(\mathbf{q})\right),
\end{aligned}
$$

where the inequality follows as $h(x) = \sqrt{x}$ is concave and hence

$$
h(y) - h(x) \leq (y-x)h'(x) = \frac{y-x}{2\sqrt{x}}.
$$

□

C.5.1 Proposition 8

We start with a lemma that established the state space collapse. Recall the definition of the imbalance process $z(k)$ in Eq (C.32).

Lemma 23 (State Space Collapse)**.** *Under the fluid pricing policy and max-weight matching policy, for all $\eta \geq 1$, $r \in \mathbb{Z}_+$ there exists a constant $\mathscr{F}_r$ (independent of η) such that*

$$
\mathbb{E}\left[\left(\langle \mathbf{1}_{2n}, \mathbf{q}^2(\infty)\rangle - \frac{z^2(\infty)}{n}\right)^{r/2}\right] \leq \mathscr{F}_r.
$$

Proof of Lemma 23. By Lemma 22, the drift of the Lyapunov function $U(\mathbf{q})$ is uniformly bounded. In the proof below, we want to show that the drift is negative outside an appropriately defined finite set. We start by calculating the drift of $V(\mathbf{q})$. By (C.13), the drift of $V^{(c)}(\mathbf{q})$ for the uniformized DTMC is

$$
\begin{aligned}
&c\,\mathbb{E}\left[\Delta V^{(c)}(\mathbf{q}) \mid \mathbf{q}(k)=\mathbf{q}\right] \\
&\quad \leq \sum_{i=1}^{n}\mu_i^* + \sum_{j=1}^{n}\lambda_j^* + 2\sum_{(i,j)\in E}\chi_{ij}^*\left[q_j^{(c)}\mathbb{1}_{\left\{q_j^{(c)}<q_{\max}^{\eta}\right\}} - \max_{j':(i,j')\in E} q_{j'}^{(c)}\right].
\end{aligned}
$$

Similarly, for $V^{(s)}(\mathbf{q})$ we will have

$$
\begin{aligned}
&c\,\mathbb{E}\left[\Delta V^{(s)}(\mathbf{q}) \mid \mathbf{q}(k)=\mathbf{q}\right] \\
&\quad \leq \sum_{i=1}^{n}\mu_i^* + \sum_{j=1}^{n}\lambda_j^* + 2\sum_{(i,j)\in E}\chi_{ij}^*\left[q_i^{(s)}\mathbb{1}_{\left\{q_i^{(s)}<q_{\max}^{\eta}\right\}} - \max_{i':(i',j)\in E} q_{i'}^{(s)}\right].
\end{aligned}
$$

Adding the two inequalities above, we have

$$
\begin{aligned}
&c\,\mathbb{E}\left[\Delta V(\mathbf{q}) \mid \mathbf{q}(k)=\mathbf{q}\right] \\
&\leq 2\sum_{i=1}^{n}\mu_i^* + 2\sum_{j=1}^{n}\lambda_j^* \\
&\qquad + 2\sum_{(i,j)\in E}\chi_{ij}^*\left[q_j^{(c)}\mathbb{1}_{\left\{q_j^{(c)}<q_{\max}^{\eta}\right\}} + q_i^{(s)}\mathbb{1}_{\left\{q_i^{(s)}<q_{\max}^{\eta}\right\}} - \max_{j':(i,j')\in E} q_{j'}^{(c)} - \max_{i':(i',j)\in E} q_{i'}^{(s)}\right] \\
&\leq 2\sum_{i=1}^{n}\mu_i^* + 2\sum_{j=1}^{n}\lambda_j^* - 2q_{\max}^{\eta}\left(\sum_{i=1}^{n}\mu_i^*\mathbb{1}_{\{q_i^{(s)}=q_{\max}^{\eta}\}} + 2\sum_{j=1}^{n}\lambda_j^*\mathbb{1}_{\{q_j^{(c)}=q_{\max}^{\eta}\}}\right) \\
&\qquad + 2\sum_{(i,j)\in E}\chi_{ij}^*\left[q_j^{(c)} + q_i^{(s)} - \max_{j':(i,j')\in E} q_{j'}^{(c)} - \max_{i':(i',j)\in E} q_{i'}^{(s)}\right] \\
&\leq 2\sum_{i=1}^{n}\mu_i^* + 2\sum_{j=1}^{n}\lambda_j^* - 2q_{\max}^{\eta}\left(\sum_{i=1}^{n}\mu_i^*\mathbb{1}_{\{q_i^{(s)}=q_{\max}^{\eta}\}} + \sum_{j=1}^{n}\lambda_j^*\mathbb{1}_{\{q_j^{(c)}=q_{\max}^{\eta}\}}\right) \\
&\qquad - \frac{2}{n}\left(\min_{i,j\in E}\chi_{ij}^*\right)U(\mathbf{q}),
\end{aligned}
$$

where the last inequality uses Lemma 21. The drift of $\Delta W_z(\mathbf{q})$ can be bounded as

$$\begin{aligned}
& c\,\mathbb{E}\left[\Delta W_z(\mathbf{q}) \mid \mathbf{q}(k)=\mathbf{q}\right] \\
&= \frac{1}{n}\mathbb{E}\left[z^2(k+1) - z^2(k) \mid \mathbf{q}(k)=\mathbf{q}\right] \\
&= \frac{1}{n}\left(\sum_{j=1}^{n}\lambda_j^* \mathbb{1}_{\{q_j^{(c)}<q_{\max}^{\eta}\}}\left((z-1)^2 - z^2\right) + \sum_{i=1}^{n}\mu_i^* \mathbb{1}_{\{q_i^{(s)}<q_{\max}^{\eta}\}}\left((z+1)^2 - z^2\right)\right) \\
&\geq -\frac{2z}{n}\left(\sum_{j=1}^{n}\lambda_j^* \mathbb{1}_{\{q_j^{(c)}<q_{\max}^{\eta}\}} - \sum_{i=1}^{n}\mu_i^* \mathbb{1}_{\{q_i^{(s)}<q_{\max}^{\eta}\}}\right) \\
&= -\frac{2z}{n}\left(\sum_{i=1}^{n}\mu_i^* \mathbb{1}_{\{q_i^{(s)}=q_{\max}^{\eta}\}} - \sum_{j=1}^{n}\lambda_j^* \mathbb{1}_{\{q_j^{(c)}=q_{\max}^{\eta}\}}\right) \\
&\geq -2q_{\max}^{\eta}\left(\sum_{i=1}^{n}\mu_i^* \mathbb{1}_{\{q_i^{(s)}=q_{\max}^{\eta}\}} + \sum_{j=1}^{n}\lambda_j^* \mathbb{1}_{\{q_j^{(c)}=q_{\max}^{\eta}\}}\right),
\end{aligned}$$

where the third equality uses the fact that $\sum_{i=1}^{n}\mu_i^* = \sum_{j=1}^{n}\lambda_j^*$, and the last inequality uses the fact that $z/n \leq q_{\max}^{\eta}$ by the definition of the fluid pricing policy.

Combining the inequalities above, by Lemma 22, for any $U(\mathbf{q}) > \frac{2n(\langle \mathbf{1}_n, \boldsymbol{\lambda}^*\rangle + \langle \mathbf{1}_n, \boldsymbol{\mu}^*\rangle)}{\min_{(i,j)\in E}\chi_{ij}^*}$, we have

$$\begin{aligned}
\mathbb{E}\left[\Delta U(\mathbf{q}) \mid \mathbf{q}(k)=\mathbf{q}\right] &\leq \frac{\mathbb{E}\left[\Delta V(\mathbf{q}) - \Delta W_z(\mathbf{q})\right]}{2U(\mathbf{q})} \\
&\leq \frac{2\left(\langle \mathbf{1}_n, \boldsymbol{\lambda}^*\rangle + \langle \mathbf{1}_n, \boldsymbol{\mu}^*\rangle\right)}{cU(\mathbf{q})} - \frac{2}{cn}\left(\min_{(i,j)\in E}\chi_{ij}^*\right) \\
&< -\frac{1}{cn}\left(\min_{(i,j)\in E}\chi_{ij}^*\right).
\end{aligned}$$

Thus, we have established that the drift of $U(\mathbf{q})$ is uniformly bounded and is negative outside a finite set. By Eryilmaz and Srikant (2012, Lemma 1), we have

$$\mathbb{E}\left[U(\mathbf{q}(\infty))^r\right] = \mathbb{E}\left[\left(\left\langle \mathbf{1}_{2n}, \mathbf{q}^2(\infty)\right\rangle - \frac{z^2(\infty)}{n}\right)^{r/2}\right] \leq \mathscr{F}_r \quad \forall\, \eta \geq 1 \quad \forall r \in \mathbb{Z}_+. \tag{C.46}$$

□

Proof of Proposition 8. In the steady state, we have $\mathbb{E}\left[\Delta W_z(\mathbf{q}(\infty))\right] = 0$. Expanding this equation, we have

$$\begin{aligned}
&\sum_{j=1}^{n} \lambda_j^* \Pr\left[q_j^{(c)}(\infty) < q_{\max}^{\eta}\right] + \sum_{i=1}^{n} \mu_i^* \Pr\left[q_i^{(s)}(\infty) < q_{\max}^{\eta}\right] \\
&= 2\,\mathbb{E}\left[z(\infty)\left(\sum_{i=1}^{n}\left(\lambda_i^* \mathbb{1}_{\{q_i^{(c)}(\infty)=q_{\max}^{\eta}\}} - \mu_i^* \mathbb{1}_{\{q_i^{(s)}(\infty)=q_{\max}^{\eta}\}}\right)\right)\right] \\
&= 2n\,\mathbb{E}\Bigg[\underbrace{\sum_{i=1}^{n}\left(q_{\perp_i}^{(s)}(\infty) - q_{\perp_i}^{(c)}(\infty) - \frac{z(\infty)}{n}\right)\left(\mu_i^* \mathbb{1}_{\{q_i^{(s)}(\infty)=q_{\max}^{\eta}\}} - \lambda_i^* \mathbb{1}_{\{q_i^{(c)}(\infty)=q_{\max}^{\eta}\}}\right)}_{(\dagger)}\Bigg] \\
&\qquad - 2n\,\mathbb{E}\Bigg[\underbrace{\sum_{i=1}^{n}\left(q_{\perp_i}^{(s)}(\infty) - q_{\perp_i}^{(c)}(\infty)\right)\left(\mu_i^* \mathbb{1}_{\{q_i^{(s)}(\infty)=q_{\max}^{\eta}\}} - \lambda_i^* \mathbb{1}_{\{q_i^{(c)}(\infty)=q_{\max}^{\eta}\}}\right)}_{(\ddagger)}\Bigg] \\
&= 2nq_{\max}^{\eta}\sum_{i=1}^{n}\left(\lambda_i^* \Pr\left[q_i^{(c)}(\infty) = q_{\max}^{\eta}\right] + \mu_i^* \Pr\left[q_i^{(s)}(\infty) = q_{\max}^{\eta}\right]\right) \\
&\qquad - 2n\,\mathbb{E}\left[\sum_{i=1}^{n}\left(q_{\perp_i}^{(s)}(\infty) - q_{\perp_i}^{(c)}(\infty)\right)\left(\mu_i^* \mathbb{1}_{\{q_i^{(s)}(\infty)=q_{\max}^{\eta}\}} - \lambda_i^* \mathbb{1}_{\{q_i^{(c)}(\infty)=q_{\max}^{\eta}\}}\right)\right]. \qquad (C.47)
\end{aligned}$$

To see why the last the last equality holds, note that

$$\begin{aligned}
q_{\perp_i}^{(s)}(\infty) - q_{\perp_i}^{(c)}(\infty) - \frac{z(\infty)}{n} &= q_i^{(s)}(\infty) - q_{\|_i}^{(s)}(\infty) - q_i^{(c)}(\infty) + q_{\|_i}^{(c)}(\infty) - \left(q_{\|_i}^{(c)}(\infty) - q_{\|_i}^{(s)}(\infty)\right) \\
&= q_i^{(s)}(\infty) - q_i^{(c)}(\infty),
\end{aligned}$$

and hence the term $(\dagger)$ can be expressed as

$$\begin{aligned}
(\dagger) &= \sum_{i=1}^{n}\left(q_i^{(s)}(\infty) - q_i^{(c)}(\infty)\right)\left(\mu_i^* \mathbb{1}_{\{q_i^{(s)}(\infty)=q_{\max}^{\eta}\}} - \lambda_i^* \mathbb{1}_{\{q_i^{(c)}(\infty)=q_{\max}^{\eta}\}}\right) \\
&= \sum_{i=1}^{n}\left(\mu_i^* q_i^{(s)}(\infty)\mathbb{1}_{\{q_i^{(s)}(\infty)=q_{\max}^{\eta}\}} + \lambda_i^* q_i^{(c)}(\infty)\mathbb{1}_{\{q_i^{(c)}(\infty)=q_{\max}^{\eta}\}}\right) \\
&= q_{\max}^{\eta}\sum_{i=1}^{n}\left(\mu_i^* \mathbb{1}_{\{q_i^{(s)}(\infty)=q_{\max}^{\eta}\}} + \lambda_i^* \mathbb{1}_{\{q_i^{(c)}(\infty)=q_{\max}^{\eta}\}}\right),
\end{aligned}$$

where the second equality above follows from Condition 3 and the fact that $q_i^{(s)}(\infty)\cdot$

$q_i^{(c)}(\infty) = 0$. Furthermore, $q_i^{(s)}(\infty) = q_{\max}^{\eta}$ implies that $q_{\perp_i}^{(s)}(\infty) \geq 0$, and $q_i^{(c)}(\infty) = q_{\max}^{\eta}$ implies that $q_{\perp_i}^{(s)}(\infty) \geq 0$. Therefore, the term $(\ddagger)$ is nonnegative. By Eq (C.47), we have

$$\begin{aligned}
&2nq_{\max}^{\eta} \sum_{i=1}^{n} \left(\lambda_i^* \Pr[q_i^{(c)}(\infty) = q_{\max}^{\eta}] + \mu_i^* \Pr[q_i^{(s)}(\infty) = q_{\max}^{\eta}] \right) \\
&\geq \sum_{j=1}^{n} \lambda_j^* \Pr\left[q_j^{(c)}(\infty) < q_{\max}^{\eta} \right] + \sum_{i=1}^{n} \mu_i^* \Pr\left[q_i^{(s)}(\infty) < q_{\max}^{\eta} \right] \\
&= \langle \mathbf{1}_n, \boldsymbol{\lambda}^* \rangle + \langle \mathbf{1}_n, \boldsymbol{\mu}^* \rangle - \sum_{i=1}^{n} \left(\lambda_i^* \Pr[q_i^{(c)}(\infty) = q_{\max}^{\eta}] + \mu_i^* \Pr[q_i^{(s)}(\infty) = q_{\max}^{\eta}] \right).
\end{aligned}$$

Rearranging the terms, we get

$$\sum_{i=1}^{n} \left(\lambda_i^* \Pr[q_i^{(c)}(\infty) = q_{\max}^{\eta}] + \mu_i^* \Pr[q_i^{(s)}(\infty) = q_{\max}^{\eta}] \right) \geq \frac{\langle \mathbf{1}_n, \boldsymbol{\lambda}^* \rangle + \langle \mathbf{1}_n, \boldsymbol{\mu}^* \rangle}{2nq_{\max}^{\eta} + 1}. \tag{C.48}$$

This proves the first part of the proposition.

Note that the proof so far applies to any matching algorithm. In the rest of the proof, we consider the max-weight algorithm specifically. By the Cauchy-Schwarz inequality, we have

$$\begin{aligned}
&\mathbb{E}\left[\sum_{i=1}^{n} \left(q_{\perp_i}^{(s)}(\infty) - q_{\perp_i}^{(c)}(\infty) \right) \left(\mu_i^* \mathbb{1}_{\{q_i^{(s)}(\infty) = q_{\max}^{\eta}\}} - \lambda_i^* \mathbb{1}_{\{q_i^{(c)}(\infty) = q_{\max}^{\eta}\}} \right) \right] \\
&\leq \sqrt{\mathbb{E}\left[\|\mathbf{q}_{\perp}^{(s)}(\infty) - \mathbf{q}_{\perp}^{(c)}(\infty)\|^2 \right]} \sqrt{\mathbb{E}\left[\sum_{i=1}^{n} \left(\mu_i^* \mathbb{1}_{\{q_i^{(s)}(\infty) = q_{\max}^{\eta}\}} - \lambda_i^* \mathbb{1}_{\{q_i^{(c)}(\infty) = q_{\max}^{\eta}\}} \right)^2 \right]} \\
&\leq \sqrt{\mathscr{F}_2} \sqrt{\sum_{i=1}^{n} \left((\mu_i^*)^2 \Pr[q_i^{(s)}(\infty) = q_{\max}^{\eta}] + (\lambda_j^*)^2 \Pr[q_j^{(c)}(\infty) = q_{\max}^{\eta}] \right)} \\
&\leq \sqrt{\mathscr{F}_2} \sqrt{\sum_{i=1}^{n} \lambda_i^* \max_{i \in N} \lambda_i^* + \sum_{i=1}^{n} \mu_i^* \max_{i \in N} \mu_i^*} \frac{1}{\sqrt{q_{\max}^{\eta}}},
\end{aligned}$$

where the second inequality follows by Lemma 23 (which requires the matching algorithm to be max-weight) and the fact that $q_i^{(s)}(\infty) \cdot q_i^{(c)}(\infty) = 0$, and the third inequality follows by (C.14), as we have

$$\sum_{i=1}^{n} \lambda_i^* \Pr[q_i^{(c)}(\infty) = q_{\max}^{\eta}] \leq \frac{1}{q_{\max}^{\eta}} \sum_{i=1}^{n} \lambda_i^*$$
$$\Rightarrow \sum_{i=1}^{n} (\lambda_i^*)^2 \Pr[q_i^{(c)}(\infty) = q_{\max}^{\eta}] \leq \frac{1}{q_{\max}^{\eta}} \sum_{i=1}^{n} \lambda_i^* \max_{i \in N} \lambda_i^*,$$

and similarly for $\Pr[q_i^{(s)}(\infty) = q_{\max}^{\eta}]$. Substituting the above inequality in Eq (C.47), we get

$$\begin{aligned} & q_{\max}^{\eta} \left(\sum_{j=1}^{n} \lambda_j^* \Pr\left[q_j^{(c)}(\infty) = q_{\max}^{\eta} \right] + \sum_{i=1}^{n} \mu_i^* \Pr\left[q_i^{(s)}(\infty) = q_{\max}^{\eta} \right] \right) \\ & \leq \frac{\langle \mathbf{1}_n, \boldsymbol{\lambda}^* \rangle + \langle \mathbf{1}_n, \boldsymbol{\mu}^* \rangle}{2n} + \sqrt{\mathscr{F}_2} \sqrt{\sum_{i=1}^{n} \lambda_i^* \max_{i \in N} \lambda_i^* + \sum_{i=1}^{n} \mu_i^* \max_{i \in N} \mu_i^*} \frac{1}{\sqrt{q_{\max}^{\eta}}}. \end{aligned} \tag{C.49}$$

Because $q_{\max}^{\eta} \to \infty$ as $\eta \to \infty$, the second term in Eq (C.49) vanishes. Combined with Eq (C.48), we obtain the second part of the proposition. $\square$

C.5.2 Proposition 9

Similar to Lemma 23, we establish the state space collapse for the two-price policy.

Lemma 24 (State Space Collapse)**.** *Under the two-price policy and max-weight matching, if $\tau_{\max}^{\eta} \sigma^{\eta} \leq \eta$, then for all $\eta \geq 1$, $r \in \mathbb{Z}_+$ there exists constants $\mathscr{T}_r$ independent of η such that*

$$\mathbb{E}\left[\left(\left\langle \mathbf{1}_{2n}, \mathbf{q}^2(\infty) \right\rangle - \frac{z^2(\infty)}{n} \right)^{r/2} \right] \leq \mathscr{T}_r.$$

Proof of Lemma 24. By Lemma 22 we already know that the drift of the Lyapunov function $U(\mathbf{q})$ is uniformly bounded. Now, we will show that the drift of $U(\mathbf{q})$ is negative outside a finite set. The drift of $V(\mathbf{q})$ for the uniformized DTMC is (by Eq (C.18))

$$
\begin{aligned}
& c\,\mathbb{E}[\Delta V^{(c)}(\mathbf{q}) \mid \mathbf{q}(k)=\mathbf{q}] \\
& = \langle \mathbf{1}_m, \boldsymbol{\lambda}^* \rangle + \langle \mathbf{1}_n, \boldsymbol{\mu}^* \rangle + 2 \sum_{i,j \in E} \chi^*_{ij} \left(q_j^{(c)} \mathbb{1}_{\left\{\max_{i':(i',j)\in E} q_{i'}^{(s)} = 0\right\}} - \max_{j':(i,j')\in E} q_{j'}^{(c)} \right) \\
& \qquad - 2\frac{\sigma^\eta}{\eta} \sum_{j=1}^{n} q_j^{(c)} \mathbb{1}_{\left\{q_j^{(c)} > \tau^\eta_{\max}\right\}} \\
& \le \langle \mathbf{1}_n, \boldsymbol{\lambda}^* \rangle + \langle \mathbf{1}_n, \boldsymbol{\mu}^* \rangle + 2n\sigma^\eta \tau^\eta_{\max}/\eta + 2 \sum_{(i,j)\in E} \chi^*_{ij} \left(q_j^{(c)} - \max_{j':(i,j')\in E} q_{j'}^{(c)} \right) \\
& \qquad - 2\frac{\sigma^\eta}{\eta} \left\langle \mathbf{1}_n, \mathbf{q}^{(c)} \right\rangle \\
& \le \langle \mathbf{1}_n, \boldsymbol{\lambda}^* \rangle + \langle \mathbf{1}_n, \boldsymbol{\mu}^* \rangle + 2n + 2 \sum_{(i,j)\in E} \chi^*_{ij} \left(q_j^{(c)} - \max_{j':(i,j')\in E} q_{j'}^{(c)} \right) - 2\frac{\sigma^\eta}{\eta} \left\langle \mathbf{1}_n, \mathbf{q}^{(c)} \right\rangle,
\end{aligned}
$$

where the last inequality uses the assumption that $\tau^\eta_{\max}\sigma^\eta \le \eta$. Similarly, by Eq (C.20), we have

$$
\begin{aligned}
c\,\mathbb{E}[\Delta V^{(s)}(\mathbf{q}) \mid \mathbf{q}(k)=\mathbf{q}] \le \langle \mathbf{1}_n, \boldsymbol{\lambda}^* \rangle + \langle \mathbf{1}_n, \boldsymbol{\mu}^* \rangle + 2n + 2 \sum_{(i,j)\in E} \chi^*_{ij} \left(q_i^{(s)} - \max_{i':(i',j)\in E} q_{i'}^{(s)} \right) \\
- 2\frac{\sigma^\eta}{\eta} \left\langle \mathbf{1}_n, \mathbf{q}^{(s)} \right\rangle.
\end{aligned}
$$

Combining the two inequalities above, since $\Delta V(\mathbf{q}) = \Delta V^{(s)}(\mathbf{q}) + \Delta V^{(c)}(\mathbf{q})$, we have

$$
\begin{aligned}
c\,\mathbb{E}[\Delta V(\mathbf{q}) \mid \mathbf{q}(k)=\mathbf{q}] \le {} & 2\langle \mathbf{1}_n, \boldsymbol{\lambda}^* \rangle + 2\langle \mathbf{1}_n, \boldsymbol{\mu}^* \rangle + 4n \\
& + 2 \sum_{(i,j)\in E} \chi^*_{ij} \left(q_i^{(s)} + q_j^{(c)} - \max_{i':(i',j)\in E} q_{i'}^{(s)} - \max_{j':(i,j')\in E} q_{j'}^{(c)} \right) \\
& - 2\frac{\sigma^\eta}{\eta} \langle \mathbf{1}_{2n}, \mathbf{q} \rangle. \qquad \text{(C.50)}
\end{aligned}
$$

Next, the drift of $W_z(\mathbf{q})$ is bounded by

$$
\begin{aligned}
& c\,\mathbb{E}\left[\Delta W_z(\mathbf{q}) \mid \mathbf{q}(k)=\mathbf{q}\right] \\
&= \frac{1}{n}\mathbb{E}\left[z^2(k+1)-z^2(k) \mid \mathbf{q}(k)=\mathbf{q}\right] \\
&= \frac{1}{n}\left(\sum_{j=1}^{n}\left(\lambda_j^* - \frac{\sigma^\eta}{\eta}\mathbb{1}_{\{q_j^{(c)}>\tau_{\max}^\eta\}}\right)\left((z-1)^2-z^2\right)\right. \\
&\qquad \left. + \sum_{i=1}^{n}\left(\mu_i^* - \frac{\sigma^\eta}{\eta}\mathbb{1}_{\{q_i^{(s)}>\tau_{\max}^\eta\}}\right)\left((z+1)^2-z^2\right)\right) \\
&= \frac{1}{n}\left(\langle \mathbf{1}_n, \boldsymbol{\lambda}^*\rangle + \langle \mathbf{1}_n, \boldsymbol{\mu}^*\rangle - \frac{\sigma^\eta}{\eta}\left(\sum_{j=1}^{n}\mathbb{1}_{\{q_j^{(c)}>\tau_{\max}^\eta\}} + \sum_{i=1}^{n}\mathbb{1}_{\{q_i^{(s)}>\tau_{\max}^\eta\}}\right)\right) \\
&\qquad - \frac{2\sigma^\eta z}{n\eta}\left(\sum_{i=1}^{n}\mathbb{1}_{\{q_i^{(s)}>\tau_{\max}^\eta\}} - \sum_{j=1}^{n}\mathbb{1}_{\{q_j^{(c)}>\tau_{\max}^\eta\}}\right) \\
&\geq \frac{1}{n}\left(\langle \mathbf{1}_n, \boldsymbol{\lambda}^*\rangle + \langle \mathbf{1}_n, \boldsymbol{\mu}^*\rangle\right) - \frac{2\sigma^\eta}{\eta} - \frac{2\sigma^\eta}{\eta}\langle \mathbf{1}_{2n}, \mathbf{q}\rangle.
\end{aligned}
\tag{C.51}
$$

where the third equality holds because $\sum_{i=1}^n \mu_i^* = \sum_{j=1}^n \lambda_j^*$ and the last inequality holds because $|z| \leq \langle \mathbf{1}_{2n}, \mathbf{q}\rangle$.

Let $B = 2(\langle \mathbf{1}_n, \boldsymbol{\lambda}^*\rangle + \langle \mathbf{1}_n, \boldsymbol{\mu}^*\rangle) + 4n + 2$. Combining Eq (C.50) and Eq (C.51), for any $U(\mathbf{q}) \geq Bn/(2\min_{(i,j)\in E}\chi_{ij}^*)$, we have

$$
\begin{aligned}
& \mathbb{E}[\Delta U(\mathbf{q}) \mid \mathbf{q}(k)=\mathbf{q}] \\
& \leq \frac{\mathbb{E}[\Delta V(\mathbf{q}) - \Delta W_z(\mathbf{q})]}{2U(\mathbf{q})} \\
& \leq \frac{1}{2cU(\mathbf{q})}\left(B + 2\sum_{(i,j)\in E}\chi_{ij}^*\left(q_i^{(s)} + q_j^{(c)} - \max_{i':(i',j)\in E} q_{i'}^{(s)} - \max_{j':(i,j')\in E} q_{j'}^{(c)}\right)\right), \\
& \leq \frac{B}{2cU(\mathbf{q})} - \frac{2}{cn}\min_{(i,j)\in E}\chi_{ij}^* \leq -\frac{1}{cn}\min_{(i,j)\in E}\chi_{ij}^*.
\end{aligned}
$$

where the first inequality is by Lemma 22 and the third inequality is by Lemma 21. This proves that the drift is negative outside a finite set. By Eryilmaz and Srikant (2012, Lemma

1), we have

$$\mathbb{E}\left[U(\mathbf{q}(\infty))^r\right] = \mathbb{E}\left[\left(\left\langle \mathbf{1}_{2n}, \mathbf{q}^2(\infty)\right\rangle - \frac{z^2(\infty)}{n}\right)^{r/2}\right] \leq \mathscr{T}_r, \quad \forall\, \eta \geq 1.$$

□

Lemma 25. *Under the two-price policy and max-weight matching policy, for all $\eta \geq 1$, we have $\mathbb{E}[z^2(\infty)] < \infty$.*

Proof of Lemma 25. We use $\sqrt{V(\mathbf{q})} = \|\mathbf{q}\|$ as the Lyapunov function for this proof. The one-step drift is

$$|\Delta\sqrt{V(\mathbf{q})}| = \big|\|\mathbf{q}(k+1)\| - \|\mathbf{q}(k)\|\big|\mathbb{1}_{\{\mathbf{q}(k)=\mathbf{q}\}} \leq \|\mathbf{q}(k+1) - \mathbf{q}(k)\|\mathbb{1}_{\{\mathbf{q}(k)=\mathbf{q}\}} \leq 1.$$

For any $\|\mathbf{q}\| \geq (2\langle \mathbf{1}_n, \boldsymbol{\lambda}^*\rangle + 2\langle \mathbf{1}_n, \boldsymbol{\mu}^*\rangle + 4n)/\sigma^\eta$, we have

$$\begin{aligned}
&\mathbb{E}\left[\Delta\sqrt{V(\mathbf{q})} \mid \mathbf{q}(k) = \mathbf{q}\right] \\
&= \mathbb{E}\left[\|\mathbf{q}(k+1)\| - \|\mathbf{q}(k)\| \big| \mathbf{q}(k) = \mathbf{q}\right] \\
&\leq \frac{1}{\|\mathbf{q}\|}\mathbb{E}\left[\|\mathbf{q}(k+1)\|^2 - \|\mathbf{q}(k)\|^2 \big| \mathbf{q}(k) = \mathbf{q}\right] \\
&= \frac{1}{c\|\mathbf{q}\|}\Big(2\langle \mathbf{1}_n, \boldsymbol{\lambda}^*\rangle + 2\langle \mathbf{1}_n, \boldsymbol{\mu}^*\rangle + 4n \\
&\qquad + 2\sum_{(i,j)\in E} \chi_{ij}^*\left(q_i^{(s)} + q_j^{(c)} - \max_{i':(i',j)\in E} q_{i'}^{(s)} - \max_{j':(i,j')\in E} q_{j'}^{(c)}\right) - 2\frac{\sigma^\eta}{\eta}\langle \mathbf{1}_{2n}, \mathbf{q}\rangle\Big) \\
&\leq \frac{1}{\|\mathbf{q}\|}\left(2\langle \mathbf{1}_n, \boldsymbol{\lambda}^*\rangle + 2\langle \mathbf{1}_n, \boldsymbol{\mu}^*\rangle + 4n - 2\frac{\sigma^\eta}{\eta}\langle \mathbf{1}_{2n}, \mathbf{q}\rangle\right) \\
&\leq \frac{2\langle \mathbf{1}_n, \boldsymbol{\lambda}^*\rangle + 2\langle \mathbf{1}_n, \boldsymbol{\mu}^*\rangle + 4n}{\|\mathbf{q}\|} - 2\frac{\sigma^\eta}{\eta} < -\frac{\sigma^\eta}{\eta},
\end{aligned}$$

where the second equality follows from Eq (C.50). So the drift is negative outside a finite set. By Eryilmaz and Srikant (2012, Lemma 1), we have $\mathbb{E}[\|\mathbf{q}\|^2] < \infty$ and hence $\mathbb{E}[z^2] < \infty$. □

Proof of Proposition 9. Consider the Lyapunov function $W_z(\mathbf{q})$. By Lemma 25, we have

$\mathbb{E}[W_z(\mathbf{q})] < \infty$. In the steady state, it holds that $\mathbb{E}[\Delta W_z(\mathbf{q}(\infty))] = 0$. By (C.51), we expand this equation to get

$$\begin{aligned}
&2\sigma^\eta\,\mathbb{E}\left[z(\infty)\left(\sum_{i=1}^{n}\mathbb{1}_{\{q_i^{(s)}(\infty)>\tau_{\max}^\eta\}}-\sum_{j=1}^{n}\mathbb{1}_{\{q_j^{(c)}(\infty)>\tau_{\max}^\eta\}}\right)\right]\\
&=\eta\,\langle\mathbf{1}_n,\boldsymbol{\lambda}^*\rangle+\eta\,\langle\mathbf{1}_n,\boldsymbol{\mu}^*\rangle\\
&\quad-\sigma^\eta\left(\sum_{j=1}^{n}\Pr\left[q_j^{(c)}(\infty)>\tau_{\max}^\eta\right]-\sum_{i=1}^{n}\Pr\left[q_i^{(s)}(\infty)>\tau_{\max}^\eta\right]\right) \qquad \text{(C.52)}\\
&\geq\eta\,\langle\mathbf{1}_n,\boldsymbol{\lambda}^*\rangle+\eta\,\langle\mathbf{1}_n,\boldsymbol{\mu}^*\rangle-\sigma^\eta n.
\end{aligned}$$

Dividing both sides by $2\sigma^\eta$, we get

$$\begin{aligned}
&\mathbb{E}\left[z(\infty)\left(\sum_{i=1}^{n}\mathbb{1}_{\{q_i^{(s)}(\infty)>\tau_{\max}^\eta\}}-\sum_{j=1}^{n}\mathbb{1}_{\{q_j^{(c)}(\infty)>\tau_{\max}^\eta\}}\right)\right]\\
&\geq\frac{\eta}{2\sigma^\eta}\left(\langle\mathbf{1}_n,\boldsymbol{\lambda}^*\rangle+\langle\mathbf{1}_n,\boldsymbol{\mu}^*\rangle\right)-\frac{n}{2}. \qquad \text{(C.53)}
\end{aligned}$$

Recall that $z(\infty)=\sum_{j=1}^{n}q_j^{(c)}(\infty)-\sum_{i=1}^{n}q_i^{(s)}(\infty)$, so

$$\begin{aligned}
z(\infty)\left(\sum_{i=1}^{n}\mathbb{1}_{\{q_i^{(s)}(\infty)>\tau_{\max}^\eta\}}-\sum_{j=1}^{n}\mathbb{1}_{\{q_j^{(c)}(\infty)>\tau_{\max}^\eta\}}\right)&\leq n\sum_{j=1}^{n}q_j^{(c)}(\infty)+n\sum_{i=1}^{n}q_i^{(s)}(\infty)\\
&=n\,\langle\mathbf{1}_{2n},\mathbf{q}(\infty)\rangle.
\end{aligned}$$

Substituting the above inequality in Eq (C.53) and dividing both sides by n, we get

$$\mathbb{E}[\langle\mathbf{1}_{2n},\mathbf{q}(\infty)\rangle]\geq\frac{\eta}{2\sigma^\eta n}\left(\langle\mathbf{1}_n,\boldsymbol{\lambda}^*\rangle+\langle\mathbf{1}_n,\boldsymbol{\mu}^*\rangle\right)-\frac{1}{2}. \qquad \text{(C.54)}$$

This proves the first part of the proposition. Note that Eq (C.54) holds for any matching algorithm.

Next, we consider the max-weight matching algorithm specifically. Using (C.52) again,

we have

$$\mathbb{E}\left[z(\infty)\left(\sum_{i=1}^{n}\mathbb{1}_{\{q_i^{(s)}(\infty)>\tau_{\max}^{\eta}\}}-\sum_{j=1}^{n}\mathbb{1}_{\{q_j^{(c)}(\infty)>\tau_{\max}^{\eta}\}}\right)\right]\le\frac{\eta}{2\sigma^{\eta}}\left(\langle\mathbf{1}_n,\boldsymbol{\lambda}^*\rangle+\langle\mathbf{1}_n,\boldsymbol{\mu}^*\rangle\right)+\frac{n}{2}.$$

Dividing both sides by n and splitting the left-hand side, we get

$$\begin{aligned}
&\mathbb{E}\left[\sum_{i=1}^{n}\left(\frac{z(\infty)}{n}-q_i^{(s)}(\infty)+q_i^{(c)}(\infty)\right)\left(\mathbb{1}_{\{q_i^{(s)}(\infty)>\tau_{\max}^{\eta}\}}-\mathbb{1}_{\{q_i^{(c)}(\infty)>\tau_{\max}^{\eta}\}}\right)\right]\\
&\qquad+\mathbb{E}\left[\sum_{i=1}^{n}\left(q_i^{(s)}(\infty)-q_i^{(c)}(\infty)\right)\left(\mathbb{1}_{\{q_i^{(s)}(\infty)>\tau_{\max}^{\eta}\}}-\mathbb{1}_{\{q_i^{(c)}(\infty)>\tau_{\max}^{\eta}\}}\right)\right]\\
&\le\frac{\eta}{2n\sigma^{\eta}}\left(\langle\mathbf{1}_n,\boldsymbol{\lambda}^*\rangle+\langle\mathbf{1}_n,\boldsymbol{\mu}^*\rangle\right)+\frac{1}{2}.
\end{aligned}\tag{C.55}$$

We bound the two expectation term separately. The first term in Eq (C.55) can be bounded by

$$\begin{aligned}
&\mathbb{E}\left[\sum_{i=1}^{n}\left(\frac{z(\infty)}{n}-q_i^{(s)}(\infty)+q_i^{(c)}(\infty)\right)\left(\mathbb{1}_{\left\{q_i^{(s)}>\tau_{\max}^{\eta}\right\}}-\mathbb{1}_{\left\{q_i^{(c)}>\tau_{\max}^{\eta}\right\}}\right)\right]\\
&\ge-\mathbb{E}\left[\sum_{i=1}^{n}\left|q_i^{(s)}(\infty)-q_i^{(c)}(\infty)-\frac{z(\infty)}{n}\right|\right]\\
&=-\mathbb{E}\left[\|\mathbf{q}_\perp^{(s)}(\infty)-\mathbf{q}_\perp^{(c)}(\infty)\|_1\right]\\
&\ge-\sqrt{n}\,\mathbb{E}\left[\|\mathbf{q}_\perp^{(s)}(\infty)-\mathbf{q}_\perp^{(c)}(\infty)\|_2\right]\\
&\ge-\sqrt{n\,\mathbb{E}\left[\|\mathbf{q}_\perp^{(s)}(\infty)-\mathbf{q}_\perp^{(c)}(\infty)\|_2^2\right]}\\
&=-\sqrt{n\,\mathbb{E}\left[\langle\mathbf{1}_{2n},\mathbf{q}^2\rangle-\frac{z^2}{n}\right]}\\
&\ge-\sqrt{n\mathscr{T}_2},
\end{aligned}$$

where the first equality uses the definition of $\mathbf{q}_\perp$, the second inequality uses the fact that $\|\mathbf{x}\|_1\le\sqrt{n}\|\mathbf{x}\|_2$ for any $\mathbf{x}\in\mathbb{R}^n$, the second inequality uses Lemma 20, and the final inequality uses Lemma 24 (which requires the matching algorithm to be max-weight) with

$r = 2$.

The second term on the left-hand side of Eq (C.55) is equal to

$$
\begin{aligned}
&\sum_{i=1}^{n} \mathbb{E}\left[\left(q_i^{(s)}(\infty) - q_i^{(c)}(\infty)\right)\left(\mathbb{1}_{\{q_i^{(s)}(\infty) > \tau_{\max}^{\eta}\}} - \mathbb{1}_{\{q_j^{(c)}(\infty) > \tau_{\max}^{\eta}\}}\right)\right] \\
&= \sum_{i=1}^{n} \mathbb{E}\left[q_i^{(s)}(\infty)\mathbb{1}_{\{q_i^{(s)}(\infty) > \tau_{\max}^{\eta}\}} + q_i^{(c)}(\infty)\mathbb{1}_{\{q_j^{(c)}(\infty) > \tau_{\max}^{\eta}\}}\right. \\
&\qquad \left. - q_i^{(s)}(\infty)\mathbb{1}_{\{q_j^{(c)}(\infty) > \tau_{\max}^{\eta}\}} - q_i^{(c)}(\infty)\mathbb{1}_{\{q_i^{(s)}(\infty) > \tau_{\max}^{\eta}\}}\right] \\
&= \sum_{i=1}^{n} \mathbb{E}\left[q_i^{(s)}(\infty)\mathbb{1}_{\{q_i^{(s)}(\infty) > \tau_{\max}^{\eta}\}} + q_i^{(c)}(\infty)\mathbb{1}_{\{q_j^{(c)}(\infty) > \tau_{\max}^{\eta}\}}\right] \\
&= \sum_{i=1}^{n} \mathbb{E}\left[q_i^{(s)}(\infty) + q_i^{(c)}(\infty)\right] - \sum_{i=1}^{n} \mathbb{E}\left[q_i^{(s)}(\infty)\mathbb{1}_{\{q_i^{(s)}(\infty) \le \tau_{\max}^{\eta}\}}\right] \\
&\qquad - \sum_{i=1}^{n} \mathbb{E}\left[q_i^{(c)}(\infty)\mathbb{1}_{\{q_i^{(c)}(\infty) \le \tau_{\max}^{\eta}\}}\right] \\
&\ge \sum_{i=1}^{n} \mathbb{E}\left[q_i^{(s)}(\infty) + q_i^{(c)}(\infty)\right] - 2n\tau_{\max}^{\eta} \\
&= \mathbb{E}\left[\langle \mathbf{1}_{2n}, \mathbf{q}(\infty)\rangle\right] - 2n\tau_{\max}^{\eta}.
\end{aligned}
$$

Note the the last two terms in the first equality are equal to zero because $q_i^{(s)}(\infty)q_i^{(c)}(\infty) = 0$ for all $i \in [n]$. Combining the two inequalities above and using Eq (C.55), we have

$$
\mathbb{E}\left[\langle \mathbf{1}_{2n}, \mathbf{q}(\infty)\rangle\right] \le \frac{\eta}{2n\sigma^{\eta}}\left(\langle \mathbf{1}_n, \boldsymbol{\lambda}^*\rangle + \langle \mathbf{1}_n, \boldsymbol{\mu}^*\rangle\right) + \frac{1}{2} + \sqrt{n\mathscr{T}_2} + 2n\tau_{\max}^{\eta}, \tag{C.56}
$$

Combining the upper bound Eq (C.56) with the lower bound Eq (C.54), under the max-weight algorithm, we have

$$
\begin{aligned}
\frac{\eta}{2n\sigma^{\eta}}\left(\langle \mathbf{1}_n, \boldsymbol{\lambda}^*\rangle + \langle \mathbf{1}_n, \boldsymbol{\mu}^*\rangle\right) - \frac{1}{2} &\le \mathbb{E}[\langle \mathbf{1}_{2n}, \mathbf{q}(\infty)\rangle] \\
&\le \frac{\eta}{2n\sigma^{\eta}}\left(\langle \mathbf{1}_n, \boldsymbol{\lambda}^*\rangle + \langle \mathbf{1}_n, \boldsymbol{\mu}^*\rangle\right) + \frac{1}{2} + \sqrt{n\mathscr{T}_2} + 2n\tau_{\max}^{\eta}.
\end{aligned}
$$

As $\eta \to \infty$, by assumption, we have $\sigma^\eta/\eta \to 0$ and $\sigma^\eta \tau^\eta_{\max}/\eta \to 0$; therefore,

$$\lim_{\eta\to\infty} \frac{\sigma^\eta}{\eta} \mathbb{E}[\langle \mathbf{1}_{2n}, \mathbf{q}(\infty)\rangle] = \frac{1}{2n} \left(\langle \mathbf{1}_n, \boldsymbol{\lambda}^*\rangle + \langle \mathbf{1}_n, \boldsymbol{\mu}^*\rangle\right).$$

□

C.5.3 Theorem 8

Lemma 26. *Under the fluid pricing policy with max-weight matching, for $q^\eta_{\max} = \sqrt{\eta/n}$, we have*

$$\limsup_{\eta\to\infty} \frac{L^\eta}{\eta^{1/2}} \leq n^{1/2} \left(\frac{\langle \mathbf{1}_n, \boldsymbol{\lambda}^*\rangle + \langle \mathbf{1}_n, \boldsymbol{\mu}^*\rangle}{2n} \max_{j\in N} F_j(\lambda_j^*) + 2 \max_{i\in N, j\in M} \{s_i^{(s)}, s_j^{(c)}\} \right).$$

Proof of Lemma 26. Recall that the loss L^η (4.8) is composed of the holding cost term and the revenue loss term. The holding cost term depends on the queue length vector $\mathbf{q}$. According to the definition of the fluid pricing policy, we have $\mathbf{q} \leq q^\eta_{\max} \mathbf{1}_{2n}$ *a.s*. Thus, it is trivially true that $\langle \mathbf{s}, \mathbb{E}[\mathbf{q}(\infty)]\rangle \leq 2 \max_{i\in N, j\in M}\{s_i^{(s)}, s_j^{(c)}\} n q^\eta_{\max}$. For the revenue loss term, we apply the upper bound in Eq (C.49) through the inequality

$$\begin{aligned}
&\eta \left\langle F(\boldsymbol{\lambda}^*), (\boldsymbol{\lambda}^* \circ \mathbb{E}[\mathbf{I}^{(c)}(q^\eta_{\max})]) \right\rangle - \eta \left\langle G(\boldsymbol{\mu}^*), (\boldsymbol{\mu}^* \circ \mathbb{E}[\mathbf{I}^{(s)}(q^\eta_{\max})] \right\rangle \\
&\leq \eta \max_{j\in N} F_j(\lambda_j^*) \left(\sum_{j=1}^n \lambda_j^* \Pr\left[q_j^{(c)}(\infty) = q^\eta_{\max}\right] + \sum_{i=1}^n \mu_i^* \Pr\left[q_i^{(s)}(\infty) = q^\eta_{\max}\right] \right).
\end{aligned}$$

Therefore, the loss L^η (4.8) is bounded by

$$\begin{aligned}
\frac{L^\eta}{\eta^{1/2}} \leq \max_{j\in N} F_j(\lambda_j^*) \Bigg(& \eta^{1/2} \frac{\langle \mathbf{1}_n, \boldsymbol{\lambda}^*\rangle + \langle \mathbf{1}_n, \boldsymbol{\mu}^*\rangle}{2n q^\eta_{\max}} \\
& + \sqrt{\mathscr{F}_2} \sqrt{\sum_{i=1}^n \lambda_i^* \max_{i\in N} \lambda_i^* + \sum_{i=1}^n \mu_i^* \max_{i\in N} \mu_i^*} \frac{\eta^{1/2}}{(q^\eta_{\max})^{3/2}} \Bigg) \\
& + 2 \max_{i\in N, j\in M} \{s_i^{(s)}, s_j^{(c)}\} \frac{n q^\eta_{\max}}{\eta^{1/2}},
\end{aligned}$$

where the constant $\mathscr{F}_2$ is defined in Lemma 23 and is independent of η. Setting $q^{\eta}_{\max} = \sqrt{\eta/n}$, we note that the term involving $\mathscr{F}_2$ converges to 0 as $\eta \to \infty$. Therefore, we get

$$\begin{aligned}\limsup_{\eta\to\infty} \frac{L^{\eta}}{\eta^{1/2}} &\leq \frac{1}{n^{1/2}} \frac{\langle \mathbf{1}_n, \boldsymbol{\lambda}^* \rangle + \langle \mathbf{1}_n, \boldsymbol{\mu}^* \rangle}{2} \max_{j\in N} F_j(\lambda_j^*) + 2n^{1/2} \max_{i\in N, j\in M} \{s_i^{(s)}, s_j^{(c)}\} \\ &= n^{1/2} \left(\frac{\langle \mathbf{1}_n, \boldsymbol{\lambda}^* \rangle + \langle \mathbf{1}_n, \boldsymbol{\mu}^* \rangle}{2n} \max_{j\in N} F_j(\lambda_j^*) + 2 \max_{i\in N, j\in M} \{s_i^{(s)}, s_j^{(c)}\} \right).\end{aligned}$$

□

Lemma 27. *Under the fluid pricing policy with randomized matching (Algorithm 10), for $q^{\eta}_{\max} = \gamma\sqrt{\eta}$, we have*

$$L^{\eta} \leq \left(\frac{\sum_{i=1}^{n} \mu_i^* + \sum_{j=1}^{m} \lambda_j^*}{2\gamma} \max_{j\in M} \{F_j(\lambda_j^*)\} + \langle \mathbf{1}_{n+m}, \mathbf{s} \rangle \gamma \right) \eta^{1/2}.$$

Proof of Lemma 27. We use $V^{(c)}(\mathbf{q}) = \left\langle \mathbf{1}_m, (\mathbf{q}^{(c)})^2 \right\rangle$ as the Lyapunov function. By Eq

(C.12), the one-step drift of $V^{(c)}(\mathbf{q})$ is bounded by

$$\begin{aligned}
\eta c\,\mathbb{E}[\Delta V^{(c)}(\mathbf{q})] &\le \sum_{i=1}^{n}\eta\mu_i^* + \sum_{j=1}^{m}\eta\lambda_j^* \\
&\quad + 2\,\mathbb{E}\left[\sum_{j=1}^{m}\eta\lambda_j^* q_j^{(c)}\left(1-\sum_{i=1}^{n}y_{ij}\right)\mathbb{1}_{\left\{q_j^{(c)}<q_{\max}^{\eta}\right\}} - \sum_{i=1}^{n}\eta\mu_i^*\sum_{j=1}^{m}q_j^{(c)}y_{ij}\right] \\
&\le \sum_{i=1}^{n}\eta\mu_i^* + \sum_{j=1}^{m}\eta\lambda_j^* + 2\,\mathbb{E}\left[\sum_{j=1}^{m}\eta\lambda_j^* q_j^{(c)}\mathbb{1}_{\left\{q_j^{(c)}<q_{\max}^{\eta}\right\}} - \sum_{i=1}^{n}\eta\mu_i^*\sum_{j=1}^{m}q_j^{(c)}y_{ij}\right] \\
&= \sum_{i=1}^{n}\eta\mu_i^* + \sum_{j=1}^{m}\eta\lambda_j^* \\
&\quad + 2\,\mathbb{E}\left[\sum_{j=1}^{m}\eta\lambda_j^* q_j^{(c)} - \sum_{i=1}^{n}\eta\mu_i^*\sum_{j=1}^{m}q_j^{(c)}y_{ij} - \sum_{j=1}^{m}\eta\lambda_j^* q_{\max}^{\eta} I_j^{(c)}(q_{\max}^{\eta})\right] \\
&\le \sum_{i=1}^{n}\eta\mu_i^* + \sum_{j=1}^{m}\eta\lambda_j^* \\
&\quad + 2\,\mathbb{E}\left[\sum_{j=1}^{m}\eta\lambda_j^* q_j^{(c)} - \sum_{i=1}^{n}\eta\mu_i^*\sum_{j=1}^{m}q_j^{(c)}\frac{\chi_{ij}^*}{\mu_i^*} - \sum_{j=1}^{m}\eta\lambda_j^* q_{\max}^{\eta} I_j^{(c)}(q_{\max}^{\eta})\right] \\
&= \sum_{i=1}^{n}\eta\mu_i^* + \sum_{j=1}^{m}\eta\lambda_j^* - 2\,\mathbb{E}\left[\sum_{j=1}^{m}\eta\lambda_j^* q_{\max}^{\eta} I_j^{(c)}(q_{\max}^{\eta})\right].
\end{aligned}$$

The third inequality above follows from the definition of the randomized matching algorithm (Algorithm 10): when $q_j^{(c)} > 0$, if a type i server arrives, the conditional probability that the server is matched to a type j customer is at least χ_{ij}^*/μ_i^*. The last equality holds because $\lambda_j^* = \sum_{i=1}^{n}\chi_{ij}^*$ for all $j \in M$ by Eq (4.6b).

In the steady state, we have $\mathbb{E}[\Delta V^{(c)}(\mathbf{q})] = 0$. Rearranging the term in the above inequality, we get

$$2\sum_{j=1}^{m}\eta\lambda_j^* q_{\max}^{\eta}\,\mathbb{E}\left[I_j^{(c)}(q_{\max}^{\eta})\right] \le \sum_{i=1}^{n}\eta\mu_i^* + \sum_{j=1}^{m}\eta\lambda_j^*,$$

which implies that

$$\sum_{j=1}^{m}\eta\lambda_j^* F_j(\lambda_j^*)\,\mathbb{E}\left[I_j^{(c)}(q_{\max}^{\eta})\right] \le \frac{\sum_{i=1}^{n}\eta\mu_i^* + \sum_{j=1}^{m}\eta\lambda_j^*}{2q_{\max}^{\eta}}\max_{j\in M}\{F_j(\lambda_j^*)\}.$$

Setting $q^\eta_{\max} = \gamma\eta^{1/2}$, we have

$$\begin{aligned}
L^\eta &= \sum_{j=1}^{m} \eta\lambda_j^* F_j(\lambda_j^*)\mathbb{E}\left[I_j^{(c)}(q^\eta_{\max})\right] + \mathbb{E}[\langle \mathbf{s}, \mathbf{q}\rangle] \\
&\le \frac{\sum_{i=1}^{n}\eta\mu_i^* + \sum_{j=1}^{m}\eta\lambda_j^*}{2q^\eta_{\max}} \max_{j\in M}\{F_j(\lambda_j^*)\} + \langle \mathbf{1}_{n+m}, \mathbf{s}\rangle\, q^\eta_{\max} \\
&= \eta^{1/2}\frac{\sum_{i=1}^{n}\mu_i^* + \sum_{j=1}^{m}\lambda_j^*}{2\gamma} \max_{j\in M}\{F_j(\lambda_j^*)\} + \langle \mathbf{1}_{n+m}, \mathbf{s}\rangle\, \gamma\eta^{1/2}.
\end{aligned}$$

□

Proof of Theorem 8. In Lemma 26, we have shown that $\limsup_{\eta\to\infty} L^\eta/\eta^{1/2} = O(n^{1/2})$ for the max-weight matching algorithm. By Lemma 27, we know that the profit loss is $O(\eta^{1/2})$ for the randomized matching policy. To complete the proof, we want to show $\liminf_{\eta\to\infty} L^\eta/\eta^{1/2} = \Omega(n)$ for the randomized matching policy.

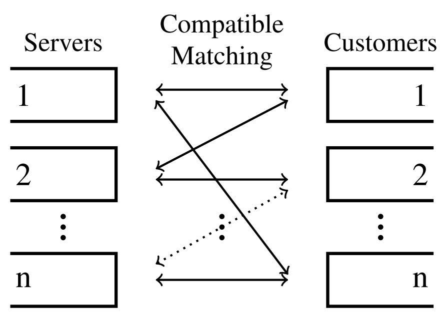


Figure C.4: The "long chain" graph used in the proof.

To this end, we consider the following instance. Consider a bipartite graph with customer type j is connected with server type j and type $j+1$ modulo n (see Figure C.4). We assume that the demand curve is $F_j(x) = 2 - 0.5x$ for all $j \in N$ and the supply curve is $G_i(x) = 0.5x$ for all $i \in N$. It is easy to verify that $\boldsymbol{\lambda}^* = \boldsymbol{\mu}^* = \mathbf{1}_n$, $\chi_{ii} = 1$ for all $i \in N$ is an optimal solution to the fluid problem. It is also easy to see that the graph and the fluid solution satisfy both Condition 2 and Condition 3. As $\chi_{ij} = 0$ for all $i \neq j$, the system behaves like n independent two-sided queues under the randomized matching algorithm. For each single-link two-sided queue (with loss L^η/n), by Eq (C.16) in Proposition 6, the loss under any fluid pricing policy is lower bounded by

$$\frac{L^{\eta}}{n} \geq \frac{1}{2q^{\eta}_{\max}+1}\eta + \min\{s^{(s)}, s^{(c)}\}\frac{q^{\eta}_{\max}(q^{\eta}_{\max}+1)}{2q^{\eta}_{\max}+1}.$$

To minimize the order of η on the right-hand side, we set $q^{\eta}_{\max} = \gamma\eta^{1/2}$. Therefore, using the AM-GM inequality $(a+b)/2 \geq \sqrt{ab}$, the total loss for this system is bounded by

$$\liminf_{\eta\to\infty} \frac{L^{\eta}}{\eta^{1/2}} \geq n\left(\frac{1}{2\gamma} + \frac{\min\{s^{(s)}, s^{(c)}\}\gamma}{2}\right) \geq n\sqrt{\min\{s^{(s)}, s^{(c)}\}}.$$

□

C.5.4 Theorem 9

Lemma 28. *Under the two-price and max-weight matching policy, for $\boldsymbol{\theta} = \mathbf{1}_n$ and $\boldsymbol{\phi} = \mathbf{1}_n$, $\sigma^{\eta} = n^{-1/3}\eta^{2/3}$ and $\tau^{\eta}_{\max} = o(\eta^{1/3})$, we have*

$$\limsup_{\eta\to\infty} \frac{L^{\eta}}{\eta^{1/3}} \leq \left(\sum_{i\in N}\left(\frac{\mu_i^* G_i''(\mu_i^*)}{2} + G_i'(\mu_i^*)\right) - \sum_{j\in N}\left(\frac{\lambda_j^* F_j''(\lambda_j^*)}{2} + F_j'(\lambda_j^*)\right)\right.$$
$$\left. + \frac{\max_{i\in N, j\in M}\{s_i^{(s)}, s_j^{(c)}\}}{2}\left(\langle \mathbf{1}_n, \boldsymbol{\lambda}^*\rangle + \langle \mathbf{1}_n, \boldsymbol{\mu}^*\rangle\right)\right) n^{-2/3} = O(n^{1/3}).$$

Proof of Lemma 28. By (C.28), the profit loss under the two-price policy is

$$\begin{aligned}
L^\eta &= \eta \sum_{j\in M} \left(F_j(\lambda_j^*)\lambda_j^* - F_j\left(\lambda_j^* - \theta_j \frac{\sigma^\eta}{\eta}\right)\left(\lambda_j^* - \theta_j \frac{\sigma^\eta}{\eta}\right)\right) \Pr[q_j^{(c)} > \tau_{\max}^\eta] \\
&\quad - \eta \sum_{i\in N} \left(G_i(\mu_i^*)\mu_i^* - G_i\left(\mu_i^* - \phi_i \frac{\sigma^\eta}{\eta}\right)\left(\mu_i^* - \phi_i \frac{\sigma^\eta}{\eta}\right)\right) \Pr[q_i^{(s)} > \tau_{\max}^\eta] \\
&\quad + \langle \mathbf{s}, \mathbb{E}[\mathbf{q}(\infty)]\rangle \\
&= \sum_{j\in M} \left(F_j'(\lambda_j^*)\lambda_j^* + F_j(\lambda_j^*)\right) \theta_j \sigma^\eta \Pr[q_j^{(c)} > \tau_{\max}^\eta] \\
&\quad - \sum_{i\in N} \left(G_i'(\mu_i^*)\mu_i^* + G_i(\mu_i^*)\right) \phi_i \sigma^\eta \Pr[q_i^{(s)} > \tau_{\max}^\eta] \\
&\quad - \left(\sum_{j\in N} \left(\frac{\lambda_j^* F_j''(\lambda_j^*)}{2} + F_j'(\lambda_j^*)\right) \theta_j^2 \Pr[q_j^{(c)} > \tau_{\max}^\eta] \right. \\
&\qquad \left. - \sum_{i\in N} \left(\frac{\mu_i^* G_i''(\mu_i^*)}{2} + G_i'(\mu_i^*)\right) \phi_i^2 \Pr[q_i^{(s)} > \tau_{\max}^\eta] \right) \frac{(\sigma^\eta)^2}{\eta} \\
&\quad + o\left(\frac{(\sigma^\eta)^2}{\eta}\right) + \langle \mathbf{s}, \mathbb{E}[\mathbf{q}(\infty)]\rangle, \qquad \text{(C.57)}
\end{aligned}$$

where the second equality uses Taylor's theorem. By Lemma 3, there exists a fluid optimal solution such that $\chi_{ij}^* > 0$ for all $(i,j) \in E$. Therefore, by Claim 1, it holds that

$$\begin{aligned}
&\sum_{j\in M} \left(F_j'(\lambda_j^*)\lambda_j^* + F_j(\lambda_j^*)\right) \theta_j \sigma^\eta \Pr[q_j^{(c)} > \tau_{\max}^\eta] \\
&\qquad - \sum_{i\in N} \left(G_i'(\mu_i^*)\mu_i^* + G_i(\mu_i^*)\right) \phi_i \sigma^\eta \Pr[q_i^{(s)} > \tau_{\max}^\eta] = 0.
\end{aligned}$$

By Eq (C.56), we have

$$\langle \mathbf{s}, \mathbb{E}[\mathbf{q}(\infty)]\rangle \le \max_{i\in N, j\in M}\{s_i^{(s)}, s_j^{(c)}\} \left(\frac{\eta}{2n\sigma^\eta} \left(\langle \mathbf{1}_n, \boldsymbol{\lambda}^*\rangle + \langle \mathbf{1}_n, \boldsymbol{\mu}^*\rangle\right) + \frac{1}{2} + \sqrt{n\mathscr{T}_2} + 2n\tau_{\max}^\eta \right).$$

Substituting the above two equations in Eq (C.57) and setting $\theta_j = \phi_i = 1$, $\sigma^\eta = n^{-1/3}\eta^{2/3}$

and $\tau_{\max}^{\eta} = o(\eta^{1/3})$, we get

$$L^{\eta} \leq \left(\sum_{i \in N} \left(\frac{\mu_i^* G_i''(\mu_i^*)}{2} + G_i'(\mu_i^*) \right) - \sum_{j \in N} \left(\frac{\lambda_j^* F_j''(\lambda_j^*)}{2} + F_j'(\lambda_j^*) \right) \right) \eta^{1/3} n^{-2/3}$$

$$+ \max_{i \in N, j \in M} \{ s_i^{(s)}, s_j^{(c)} \} \left(\frac{\eta^{1/3}}{2n^{2/3}} \left(\langle \mathbf{1}_n, \boldsymbol{\lambda}^* \rangle + \langle \mathbf{1}_n, \boldsymbol{\mu}^* \rangle \right) + \frac{1}{2} + \sqrt{n \mathscr{T}_2} + o(\eta^{1/3}) \right)$$

$$+ o(\eta^{1/3}).$$

Thus, we have

$$\limsup_{\eta \to \infty} \frac{L^{\eta}}{\eta^{1/3}} \leq \left(\sum_{i \in N} \left(\frac{\mu_i^* G_i''(\mu_i^*)}{2} + G_i'(\mu_i^*) \right) - \sum_{j \in N} \left(\frac{\lambda_j^* F_j''(\lambda_j^*)}{2} + F_j'(\lambda_j^*) \right) \right.$$

$$\left. + \frac{\max_{i \in N, j \in M} \{ s_i^{(s)}, s_j^{(c)} \}}{2} \left(\langle \mathbf{1}_n, \boldsymbol{\lambda}^* \rangle + \langle \mathbf{1}_n, \boldsymbol{\mu}^* \rangle \right) \right) n^{-2/3}$$

$$= O(n^{1/3}).$$

□

Lemma 29. *Under the two-price and randomized matching policy (Algorithm 10), for $\sigma^{\eta} = \eta^{2/3}$ and $\tau_{\max}^{\eta} = \gamma \eta^{1/3}$, we have*

$$L^{\eta} = O(\eta^{1/3}).$$

Proof of Lemma 29. We use $V^{(c)}(\mathbf{q}) = \left\langle \mathbf{1}_m, (\mathbf{q}^{(c)})^2 \right\rangle$ as the Lyapunov function. By Eq

(C.12), the one-step drift of $V^{(c)}(\mathbf{q})$ is bounded by

$$
\begin{aligned}
\eta c\,\mathbb{E}\left[\Delta V^{(c)}(\mathbf{q})\right] = \mathbb{E}\Bigg[&\sum_{j=1}^{m}\left(\eta\lambda_j^* - \theta_j\sigma^\eta \mathbb{1}_{\left\{q_j^{(c)}>\tau_{\max}^\eta\right\}}\right)\left(1-\sum_{i=1}^{n} y_{ij}\right)\left(1+2q_j^{(c)}\right) \\
&+\sum_{i=1}^{n}\left(\eta\mu_i^* - \phi_i\sigma^\eta \mathbb{1}_{\left\{q_i^{(s)}>\tau_{\max}^\eta\right\}}\right)\sum_{j=1}^{m} y_{ij}\left(1-2q_j^{(c)}\right)\Bigg] \\
\le \mathbb{E}\Bigg[&\sum_{j=1}^{m}\left(\eta\lambda_j^* - \theta_j\sigma^\eta \mathbb{1}_{\left\{q_j^{(c)}>\tau_{\max}^\eta\right\}}\right)\left(1+2q_j^{(c)}\right) \\
&+\sum_{i=1}^{n}\left(\eta\mu_i^* - \phi_i\sigma^\eta \mathbb{1}_{\left\{q_i^{(s)}>\tau_{\max}^\eta\right\}}\right)\sum_{j=1}^{m} y_{ij}\left(1-2q_j^{(c)}\right)\Bigg] \\
\le \eta\left(\langle \mathbf{1}_n,\boldsymbol{\mu}^*\rangle + \langle \mathbf{1}_m,\boldsymbol{\lambda}^*\rangle\right) &+ 2\,\mathbb{E}\Bigg[\eta\left(\sum_{j=1}^{m}\lambda_j^* q_j^{(c)} - \sum_{i=1}^{n}\mu_i^*\sum_{j=1}^{m} y_{ij}q_j^{(c)}\right) \\
&-2\sigma^\eta\sum_{j=1}^{m}\theta_j q_j^{(c)}\mathbb{1}_{\left\{q_j^{(c)}>\tau_{\max}^\eta\right\}} + 2\sigma^\eta\sum_{i=1}^{n}\phi_i\mathbb{1}_{\left\{q_i^{(s)}>\tau_{\max}^\eta\right\}}\sum_{j=1}^{m} y_{ij}q_j^{(c)}\Bigg] \\
\le \eta\left(\langle \mathbf{1}_n,\boldsymbol{\mu}^*\rangle + \langle \mathbf{1}_m,\boldsymbol{\lambda}^*\rangle\right) &+ 2\,\mathbb{E}\Bigg[\eta\left(\sum_{j=1}^{m}\lambda_j^* q_j^{(c)} - \sum_{i=1}^{n}\mu_i^*\sum_{j=1}^{m} \chi_{ij}^* q_j^{(c)}\right) \\
&-2\sigma^\eta\sum_{j=1}^{m}\theta_j q_j^{(c)}\mathbb{1}_{\left\{q_j^{(c)}>\tau_{\max}^\eta\right\}} + 2\sigma^\eta\sum_{i=1}^{n}\phi_i\mathbb{1}_{\left\{q_i^{(s)}>\tau_{\max}^\eta\right\}}\sum_{j=1}^{m} \chi_{ij}^* q_j^{(c)}\Bigg] \\
=\eta\left(\langle \mathbf{1}_n,\boldsymbol{\mu}^*\rangle + \langle \mathbf{1}_m,\boldsymbol{\lambda}^*\rangle\right)& \\
&-2\sigma^\eta\,\mathbb{E}\left[\sum_{j=1}^{m}\theta_j q_j^{(c)}\mathbb{1}_{\left\{q_j^{(c)}>\tau_{\max}^\eta\right\}} - \sum_{i=1}^{n}\phi_i\mathbb{1}_{\left\{q_i^{(s)}>\tau_{\max}^\eta\right\}}\sum_{j=1}^{m}\frac{\chi_{ij}^*}{\mu_i^*}q_j^{(c)}\right] \\
=\eta\left(\langle \mathbf{1}_n,\boldsymbol{\mu}^*\rangle + \langle \mathbf{1}_m,\boldsymbol{\lambda}^*\rangle\right)& - 2\sigma^\eta\,\mathbb{E}\left[\sum_{j=1}^{m}\theta_j q_j^{(c)}\mathbb{1}_{\left\{q_j^{(c)}>\tau_{\max}^\eta\right\}}\right],
\end{aligned}
$$

where the second inequality uses $\sum_{j=1}^{m}\lambda_j^* = \sum_{i=1}^{n}\mu_i^*$. The third inequality follows by the definition of Algorithm 10: when $q_j^{(c)} > 0$, if a type i server arrives, the conditional probability that the server is matched to a type j customer is at least χ_{ij}^*/μ_i^*. The second equality uses the fact that $\lambda_j^* = \sum_{i=1}^{n}\chi_{ij}^*$ for all $j \in M$ by Eq (4.6b). The last equality follows as for all $(i,j) \in E$ and $\mathbf{q} \in S$, we have $q_i^{(s)}q_j^{(c)} = 0$.

Similarly, we can calculate the drift of $V^{(s)}(\mathbf{q})$ to get

$$\eta c\,\mathbb{E}[\Delta V^{(s)}(\mathbf{q})] \leq \eta\left(\langle \mathbf{1}_n, \boldsymbol{\mu}^*\rangle + \langle \mathbf{1}_m, \boldsymbol{\lambda}^*\rangle\right) - 2\sigma^\eta\,\mathbb{E}\left[\sum_{i=1}^{n} \phi_i q_i^{(s)} \mathbb{1}_{\left\{q_i^{(s)} > \tau_{\max}^\eta\right\}}\right].$$

Combining the two inequalities above, since $V(\mathbf{q}) = V^{(c)}(\mathbf{q}) + V^{(s)}(\mathbf{q})$, we have

$$\begin{aligned}\eta c\,\mathbb{E}[\Delta V(\mathbf{q})] \leq 2\eta\left(\langle \mathbf{1}_n, \boldsymbol{\mu}^*\rangle + \langle \mathbf{1}_m, \boldsymbol{\lambda}^*\rangle\right) - 2\sigma^\eta \sum_{j=1}^{m} \theta_j q_j^{(c)} \mathbb{1}_{\left\{q_j^{(c)} > \tau_{\max}^\eta\right\}} \\ - 2\sigma^\eta \sum_{i=1}^{n} \phi_i q_i^{(s)} \mathbb{1}_{\left\{q_i^{(s)} > \tau_{\max}^\eta\right\}}.\end{aligned}$$

Thus, the drift is negative outside the finite set $\mathscr{B}$ defined as

$$\begin{aligned}\mathscr{B} \triangleq \Bigg\{\mathbf{q} : q_j^{(c)} \leq \max\left\{\tau_{\max}^\eta, \frac{2\eta}{\theta_j \sigma^\eta}\left(\langle \mathbf{1}_n, \boldsymbol{\mu}^*\rangle + \langle \mathbf{1}_m, \boldsymbol{\lambda}^*\rangle\right)\right\}, \\ q_i^{(s)} \leq \max\left\{\tau_{\max}^\eta, \frac{2\eta}{\phi_i \sigma^\eta}\left(\langle \mathbf{1}_n, \boldsymbol{\mu}^*\rangle + \langle \mathbf{1}_m, \boldsymbol{\lambda}^*\rangle\right)\right\}\Bigg\}.\end{aligned}$$

By the Foster-Lyapunov theorem, the uniformized DTMC is positive recurrent. Moreover, using the moment bound theorem as a corollay of the Foster-Lyapunov theorem (Hajek, 2006), we get

$$\begin{aligned}\mathbb{E}\left[\left\langle \boldsymbol{\theta}, \mathbf{q}^{(c)}\right\rangle\right] + \mathbb{E}\left[\left\langle \boldsymbol{\phi}, \mathbf{q}^{(s)}\right\rangle\right] \leq \tau_{\max}^\eta \left(\sum_{j=1}^{m} \theta_j \Pr[q_j^{(c)} > \tau_{\max}^\eta] + \sum_{i=1}^{n} \phi_i \Pr[q_i^{(s)} > \tau_{\max}^\eta]\right) \\ + \frac{\eta}{\sigma^\eta}\left(\langle \mathbf{1}_n, \boldsymbol{\mu}^*\rangle + \langle \mathbf{1}_m, \boldsymbol{\lambda}^*\rangle\right).\end{aligned}$$

Given the above bound on the expected queue length, we can bound the profit loss to be $O(\eta^{1/3})$ using the identical proof in Theorem 6, so we omit the details. □

Proof of Theorem 9. By Lemma 28 we have the required bound for profit loss under max-weight matching. By Lemma 29, we know that the profit loss is $O(\eta^{1/3})$ under randomized matching policy. To complete the proof, we will show that $\liminf_{\eta\to\infty} L^\eta/\eta^{1/3} = \Omega(n)$ for the two-price randomized matching algorithm.

We use the same graph and fluid solution as in the proof of Theorem 9 (see Figure C.4) and show that the limiting scaled profit loss is of the order n. Note that as $\chi_{ij} = 0$ for all $i \neq j$, the system behaves like n independent two-sided queue under the randomized matching algorithm. For each single-link two-sided queue under any two-price pricing policy (which has loss L^η/n), by (C.39), the loss for for any $\sigma^\eta = \gamma\eta^{2/3}$ is bounded by

$$\frac{L^\eta}{n} \geq \left(\min_{i\in N, j\in M} \{A_j, B_i\} \frac{\gamma K_2}{4} + \min_{i\in N, j\in M} \{s_i^{(s)}, s_j^{(c)}\} \max_{i\in N, j\in M} \left\{ \frac{1}{2}, \frac{\lambda_j^*}{2(\theta_j+\phi_i)\gamma} - 0.5 \right\} \right) \eta^{1/3} \tag{C.58}$$

$$\geq \frac{\min_{i\in N, j\in M}\{s_i^{(s)}, s_j^{(c)}\}}{2} \eta^{1/3}.$$

(Recall that A_j and B_i represent the second derivative of $-F(\lambda_j)\lambda_j$ and $G(\mu_i)\mu_i$, respectively.) Therefore, we have

$$\liminf_{\eta\to\infty} \frac{L^\eta}{\eta^{1/3}} \geq \frac{\min_{i\in N, j\in M}\{s_i^{(s)}, s_j^{(c)}\}}{2} n.$$

□

C.5.5 Relaxing Condition 3

In this section, we show that Condition 3 can be relaxed without loss of generality. Suppose we are given a graph $G(N_s \cup N_c, E)$ and the fluid solution $(\boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)$ such that the CRP condition is satisfied but Condition 3 is violated. We will construct another graph $G'(N_s' \cup N_c', E')$ such that Condition 3 is satisfied and use max-weight matching and fluid price policy/two-price policy for this new graph. Then, we will conduct a similar analysis as in Theorem 8 and Theorem 9 to get the limiting bound on the scaled loss in profit.

Fluid price policy and max-weight matching policy.

The idea of the graph construction is to split each customer/server type into multiple types so that we can eventually achieve the same number of customer and server types. We assume that there exists a $\delta > 0$ and such that the fluid solution $(\boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)$ are integral multiple of δ. (In practice, if the model parameters are rational numbers, this can always be achieved.) In addition, let $\lambda_{\max}$ be an upper bound of the entries in $(\boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)$.

Specifically, for each customer type $j \in M$, we split arrivals into λ_j^*/δ separate queues. Let $q_{jl}^{(c)}$ for all $l \in [\lambda_j^*/\delta]$ denote the queue length of the l^{th} replication. Suppose the arrival rate for (original) type j customer is $\lambda_j^\eta(\mathbf{q}) = \eta\lambda_j^* - \delta\sum_{i=1}^{\lambda_j^*/\delta} \mathbb{1}_{\{q_{jl}^{(c)} = q_{\max}^\eta\}}$. Then, the arrival rate of each split queue is $\lambda_{jl}^\eta(\mathbf{q}) = \eta\delta\mathbb{1}_{\{q_{jl}^{(c)} < q_{\max}^\eta\}}$. We split the server types similarly.

This leads to a modified graph $G'(N_s' \cup N_c', E')$ with $N_c = \{jl : l \in [\lambda_j/\delta], j \in M\}$, $|N_c'| = \langle \mathbf{1}_m, \boldsymbol{\lambda}^* \rangle/\delta$ and $N_s = \{im : m \in [\mu_i/\delta], i \in N\}$, $|N_s'| = \langle \mathbf{1}_n, \boldsymbol{\mu}^* \rangle/\delta$. The edges E' of the modified graph satisfy the following: $(i,j) \in E$ if and only for all $l \in [\lambda_j/\delta]$ and for all $m \in [\mu_i/\delta]$, we have $(im, jl) \in E'$. We continue to use the max-weight matching algorithm under the modified graph $G'(N_s' \cup N_c', E')$.

Next, we show that Condition 3 holds for the modified graph. Since $\langle \mathbf{1}_m, \boldsymbol{\lambda}^* \rangle = \langle \mathbf{1}_n, \boldsymbol{\mu}^* \rangle$, we have $|N_c'| = |N_s'| = \langle \mathbf{1}_m, \boldsymbol{\lambda}^* \rangle/\delta \le n\lambda_{\max}/\delta$. In addition, for any $J \subsetneq N_c'$, we have

$$\begin{aligned}
\delta|J| = \sum_{jl \in J} \delta \le \sum_{j \in M : \exists l, jl \in J} \lambda_j^* < \sum_{i : \exists j, \exists l, (i,j) \in E, jl \in J} \mu_i^* = \sum_{i \in N : \exists jl \in J, (i,j) \in E} \sum_{m \in [\mu_i^*/\delta]} \delta \\
= \sum_{im \in N_c' : \exists jl \in J, (im, jl) \in E'} \delta = \delta|N(J)|,
\end{aligned} \tag{C.59}$$

where the second inequality follows from the CRP condition and the third equality follows as

$$\{im : i \in M, m \in [\mu_i^*/\delta], \exists jl \in J, (i,j) \in E\} \iff \{im \in N_c' : \exists jl \in J, (im, jl) \in E'\}.$$

Eq (C.59) is Hall's condition for all the customer types. We can similarly verify Hall's condition for any subset of servers. This implies that there exists a perfect matching in $G'(N_s' \cup N_c', E')$. Thus, Condition 3 is satisfied for this modified graph.

Applying Proposition 8 to the modified graph, we have

$$\begin{aligned}
&\lim_{\eta\to\infty} q_{\max}^{\eta} \left(\sum_{j=1}^{m} \sum_{l=1}^{\lambda_j^*/\delta} \lambda_j^* \Pr\left[q_{jl}^{(c)} = q_{\max}^{\eta} \right] + \sum_{i=1}^{n} \sum_{m=1}^{\mu_i^*/\delta} \mu_i^* \Pr\left[q_{im}^{(s)} = q_{\max}^{\eta} \right] \right) \\
&\quad = \frac{\langle \mathbf{1}_n, \boldsymbol{\lambda}^* \rangle + \langle \mathbf{1}_n, \boldsymbol{\mu}^* \rangle}{2n}.
\end{aligned}$$

In addition, it is trivially true that

$$\mathbb{E}[\langle \mathbf{1}_{n+m}, \mathbf{q} \rangle] \leq \frac{\langle \mathbf{1}_m, \boldsymbol{\mu}^* \rangle}{\delta} q_{\max}^{\eta} \leq n \frac{\lambda_{\max}}{\delta} q_{\max}^{\eta}.$$

The above two results together with $q_{\max}^{\eta} = n^{-1/2}\eta^{1/2}$ will produce $\limsup_{\eta\to\infty} L^{\eta}/\sqrt{\eta} = O(\sqrt{n})$.

Two-price policy and max-weight matching policy

As in previous subsection, we split customer and server types to satisfy Condition 3. For customer type j, let $q_{jl}^{(c)}$ for all $l \in \left[\lambda_j^*/\delta \right]$ denote the queue length for the l^{th} replication. The arrival rate for the l^{th} replication is $\lambda_{jl}^{\eta}(\mathbf{q}) = \eta\delta - \sigma^{\eta} \mathbb{1}_{\left\{ q_{jl}^{(c)} > \tau_{\max}^{\eta} \right\}}$. Similarly, we replicate the servers as well. We apply the two-price max-weight matching policy for the modified graph $G'(N_s' \cup N_c', E')$. By Proposition 9, we have

$$\lim_{\eta\to\infty} \frac{\sigma^{\eta}}{\eta} \mathbb{E}[\langle \mathbf{1}_{2n}, \mathbf{q} \rangle] = \delta \frac{\langle \mathbf{1}_m, \boldsymbol{\lambda}^* \rangle + \langle \mathbf{1}_n, \boldsymbol{\mu}^* \rangle}{2 \langle \mathbf{1}_m, \boldsymbol{\lambda}^* \rangle}.$$

Next, we analyze the profit loss similarly as in the proof of Theorem 9. By using the definition of the pricing policy and Taylor's theorem, we get

$$
\begin{aligned}
L^\eta = \eta\Bigg(& \sum_{j=1}^{m} F_j(\lambda_j^*)\lambda_j^* \\
& - \mathbb{E}\left[\sum_{j=1}^{m}\left(\lambda_j^* - \frac{\sigma^\eta}{\eta}\sum_{l=1}^{\lambda_j^*/\delta} \mathbb{1}_{\{q_{jl} > \tau_{\max}^\eta\}}\right) F_j\left(\lambda_j^* - \frac{\sigma^\eta}{\eta}\sum_{l=1}^{\lambda_j^*/\delta} \mathbb{1}_{\{q_{jl} > \tau_{\max}^\eta\}}\right)\right]\Bigg) \\
- \eta\Bigg(& \sum_{i=1}^{n} G_i(\mu_i^*)\mu_i^* \\
& - \mathbb{E}\left[\sum_{i=1}^{n}\left(\mu_i^* - \frac{\sigma^\eta}{\eta}\sum_{m=1}^{\mu_i^*/\delta} \mathbb{1}_{\{q_{im} > \tau_{\max}^\eta\}}\right) G_i\left(\mu_i^* - \frac{\sigma^\eta}{\eta}\sum_{m=1}^{\mu_i^*/\delta} \mathbb{1}_{\{q_{im} > \tau_{\max}^\eta\}}\right)\right]\Bigg) \\
= \sigma^\eta\Bigg(& \sum_{j=1}^{m}\left(\lambda_j^* F_j'(\lambda_j^*) + F_j(\lambda_j^*)\right)\mathbb{E}\left[\sum_{l=1}^{\lambda_j^*/\delta} \mathbb{1}_{\{q_{jl} > \tau_{\max}^\eta\}}\right] \\
& - \sum_{i=1}^{m}\left(\mu_i^* G_i'(\mu_i^*) + G_i(\mu_i^*)\right)\mathbb{E}\left[\sum_{m=1}^{\mu_i^*/\delta} \mathbb{1}_{\{q_{im} > \tau_{\max}^\eta\}}\right]\Bigg) \\
& - \frac{(\sigma^\eta)^2}{\eta}\sum_{j=1}^{m}\left(\lambda_j^* F_j''(\lambda_j^*)\,\mathbb{E}\left[\sum_{l=1}^{\lambda_j^*/\delta} \mathbb{1}_{\{q_{jl} > \tau_{\max}^\eta\}}\right] + F_j'(\lambda_j^*)\,\mathbb{E}\left[\left(\sum_{l=1}^{\lambda_j^*/\delta} \mathbb{1}_{\{q_{jl} > \tau_{\max}^\eta\}}\right)^2\right]\right) \\
& + \frac{(\sigma^\eta)^2}{\eta}\sum_{i=1}^{n}\left(\mu_i^* G_i''(\mu_i^*)\,\mathbb{E}\left[\sum_{m=1}^{\mu_i^*/\delta} \mathbb{1}_{\{q_{im} > \tau_{\max}^\eta\}}\right] + G_i'(\mu_i^*)\,\mathbb{E}\left[\left(\sum_{m=1}^{\mu_i^*/\delta} \mathbb{1}_{\{q_{im} > \tau_{\max}^\eta\}}\right)^2\right]\right) \\
& + o(\eta^{1/3}) \\
= & -\frac{(\sigma^\eta)^2}{\eta}\sum_{j=1}^{m}\left(\lambda_j^* F_j''(\lambda_j^*)\,\mathbb{E}\left[\sum_{l=1}^{\lambda_j^*/\delta} \mathbb{1}_{\{q_{jl} > \tau_{\max}^\eta\}}\right] + F_j'(\lambda_j^*)\,\mathbb{E}\left[\left(\sum_{l=1}^{\lambda_j^*/\delta} \mathbb{1}_{\{q_{jl} > \tau_{\max}^\eta\}}\right)^2\right]\right) \\
& + \frac{(\sigma^\eta)^2}{\eta}\sum_{i=1}^{n}\left(\mu_i^* G_i''(\mu_i^*)\,\mathbb{E}\left[\sum_{m=1}^{\mu_i^*/\delta} \mathbb{1}_{\{q_{im} > \tau_{\max}^\eta\}}\right] + G_i'(\mu_i^*)\,\mathbb{E}\left[\left(\sum_{m=1}^{\mu_i^*/\delta} \mathbb{1}_{\{q_{im} > \tau_{\max}^\eta\}}\right)^2\right]\right) \\
& + o(\eta^{1/3}) \\
\le & \frac{(\sigma^\eta)^2\lambda_{\max}}{\eta\delta}\left(\sum_{i=1}^{n}\mu_i^* G_i''(\mu_i^*) + G_i'(\mu_i^*) - \sum_{j=1}^{m}\lambda_j^* F_j''(\lambda_j^*) + F_j'(\lambda_j^*)\right) + o(\eta^{1/3}),
\end{aligned}
$$

where the third equality holds by claim 1: as the CRP condition is satisfied, there exists a fluid solution such that $\chi^*_{ij} > 0$ for all $(i,j) \in E$. Setting $\sigma^\eta = \eta^{2/3} n^{-1/3}$, we get

$$\limsup_{\eta\to\infty} \frac{L^\eta}{\eta^{1/3}} \le n^{1/3} \Bigg(\lambda_{\max} \frac{\sum_{i=1}^n \mu_i^* G_i''(\mu_i^*) + G_i'(\mu_i^*) - \sum_{j=1}^m \lambda_j^* F_j''(\lambda_j^*) + F_j'(\lambda_j^*)}{n\delta} + \delta \max_{i\in N, j\in M} \{s_i^{(s)}, s_j^{(c)}\} \Bigg).$$

REFERENCES

Adan, Ivo and Gideon Weiss (2012). "Exact FCFS matching rates for two infinite multitype sequences". In: *Operations Research* 60.2, pp. 475–489.

Adelman, Daniel (2007). "Dynamic bid prices in revenue management". In: *Operations Research* 55.4, pp. 647–661.

Ahuja, Ravindra K., Thomas L. Magnanti, and James B. Orlin (1993). "Network Flows: Theory". In: *Algorithms, and Applications* 526.

Akbarpour, Mohammad, Shengwu Li, and Shayan Oveis Gharan (2020). "Thickness and information in dynamic matching markets". In: *Journal of Political Economy* 128.3, pp. 783–815.

Anderson, Ross, Itai Ashlagi, David Gamarnik, and Yash Kanoria (2017). "Efficient dynamic barter exchange". In: *Operations Research* 65.6, pp. 1446–1459.

Arlotto, Alessandro and Itai Gurvich (2019). "Uniformly Bounded Regret in the Multisecretary Problem". In: *Stochastic Systems* 9.3, pp. 231–260.

Arlotto, Alessandro and Xinchang Xie (2020). "Logarithmic Regret in the Dynamic and Stochastic Knapsack Problem with Equal Rewards". In: *Stochastic Systems* 10.2, pp. 170–191.

Armony, Mor, Nahum Shimkin, and Ward Whitt (2009). "The impact of delay announcements in many-server queues with abandonment". In: *Operations Research* 57.1, pp. 66–81.

Assad, Arjang A. (1978). "Multicommodity Network Flows—A Survey". In: *Networks* 8.1, pp. 37–91.

Atar, Rami and Martin I. Reiman (2012). "Asymptotically Optimal Dynamic Pricing for Network Revenue Management". In: *Stochastic Systems* 2.2, pp. 232–276.

Babaioff, Moshe, Nicole Immorlica, David Kempe, and Robert Kleinberg (2007). "A Knapsack Secretary Problem with Applications". In: *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques*. Ed. by Moses Charikar, Klaus Jansen, Omer Reingold, and José D. P. Rolim. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 16–28. ISBN: 978-3-540-74208-1.

Banerjee, Siddhartha, Daniel Freund, and Thodoris Lykouris (2017). "Pricing and Optimization in Shared Vehicle Systems: An Approximation Framework". In: *Proceed-*

ings of the 2017 ACM Conference on Economics and Computation. Cambridge, Massachusetts, USA, pp. 517–517.

Banerjee, Siddhartha, Ramesh Johari, and Carlos Riquelme (2016). “Dynamic pricing in ridesharing platforms”. In: *ACM SIGecom Exchanges* 15.1, pp. 65–70.

Banerjee, Siddhartha, Yash Kanoria, and Pengyu Qian (2018). “State dependent control of closed queueing networks”. In: *ACM SIGMETRICS Performance Evaluation Review* 46.1, pp. 2–4.

Berbeglia, Gerardo (2018). “The generalized stochastic preference choice model”. In: *arXiv preprint arXiv:1803.04244*.

Bertsekas, Dimitri P. (2005). *Dynamic Programming and Optimal Control, 3rd Edition*. Athena Scientific. ISBN: 1886529264.

Bertsekas, Dimitri P (2007). *Dynamic Programming and Optimal Control*. 4th. Vol. II. Athena Scientific.

Besbes, Omar, Francisco Castro, and Ilan Lobel (2018). *Surge pricing and its spatial supply response*. Columbia Business School Research Paper.

Besbes, Omar and Assaf Zeevi (2012). “Blind Network Revenue Management”. In: *Operations Research* 60.6, pp. 1537–1550.

Bodea, Tudor, Mark Ferguson, and Laurie Garrow (2009). “Data Set—Choice-Based Revenue Management: Data from a Major Hotel Chain”. In: *Manufacturing & Service Operations Management* 11.2, pp. 356–361.

Boucheron, Stéphane, Gábor Lugosi, and Pascal Massart (2013). *Concentration Inequalities: A Nonasymptotic Theory of Independence*. Oxford: Oxford University Press. 496 pp. ISBN: 978-0-19-953525-5.

Bronnenberg, Bart J., Michael W. Kruger, and Carl F. Mela (2008). “Database Paper—The IRI Marketing Data Set”. In: *Marketing Science* 27.4, pp. 745–748.

Bronnenberg, Bart J. and Carl F. Mela (2004). “Market Roll-Out and Retailer Adoption for New Brands”. In: *Marketing Science* 23, pp. 500–518.

Caldentey, René, Edward H Kaplan, and Gideon Weiss (2009). “FCFS infinite bipartite matching of servers and customers”. In: *Advances in Applied Probability* 41.3, pp. 695–730.

Cao, Yufeng, Anton Kleywegt, and He Wang (2019). “Network revenue management under a spiked multinomial logit choice model”. In: *Available at SSRN*.

Chen, Hong and Murray Z Frank (2001). “State dependent pricing with a queue”. In: *IIE Transactions* 33.10, pp. 847–860.

Chen, Lijian and Tito Homem-de-Mello (2010). “Re-Solving Stochastic Programming Models for Airline Revenue Management”. In: *Annals of Operations Research* 177.1, pp. 91–114.

Chen, Ningyuan, Guillermo Gallego, and Zhuodong Tang (2021). “The Use of Binary Choice Forests to Model and Estimate Discrete Choices”. In:

Chen, Yi-Chun and Velibor V Mišić (2019). “Decision forest: A nonparametric approach to modeling irrational choice”. In: *arXiv preprint arXiv:1904.11532*.

Cooper, William L. (2002). “Asymptotic Behavior of an Allocation Policy for Revenue Management”. In: *Operations Research* 50.4, pp. 720–727.

Crainic, Teodor Gabriel (2000). “Service Network Design in Freight Transportation”. In: *European Journal of Operational Research* 122.2, pp. 272–288.

Dijkstra, Edsger W. (Dec. 1, 1959). “A Note on Two Problems in Connexion with Graphs”. In: *Numerische Mathematik* 1.1, pp. 269–271.

Domencich, Thomas A. and Daniel McFadden (1975). *Urban Travel Demand: A Behavioral Analysis : a Charles River Associates Research Study*. Charles River Associates. Research Studies. Charles River Associates. Research Studies. Contributions to economic analysis, 93 v. 93. North-Holland Publishing Company. ISBN: 9780444108302.

Echenique, F. and K. Saito (2018). “General Luce model”. In: *Economic Theory*, pp. 1–16.

Echenique, Federico, Kota Saito, and Gerelt Tserenjigmid (2018). “The perception-adjusted Luce model”. In: *Mathematical Social Sciences* 93, pp. 67–76.

Eryilmaz, Atilla and Rayadurgam Srikant (2012). “Asymptotically tight steady-state queue length bounds implied by drift conditions”. In: *Queueing Systems* 72.3-4, pp. 311–359.

Ferreira, Kris Johnson, David Simchi-Levi, and He Wang (2018). “Online Network Revenue Management Using Thompson Sampling”. In: *Operations Research* 66.6, pp. 1586–1602.

Ford, Lester R. and Delbert R. Fulkerson (Oct. 1958). “A Suggested Computation for Maximal Multi-Commodity Network Flows”. In: *Management Science* 5.1, 97–101.

Frank, Marguerite and Philip Wolfe (1956). “An Algorithm for Quadratic Programming”. In: *Naval Research Logistics Quarterly* 3.1-2, pp. 95–110.

Freedman, David A. (1975). “On Tail Probabilities for Martingales”. In: *The Annals of Probability* 3.1, pp. 100–118.

Gallego, Guillermo, Richard Ratliff, and Sergey Shebalov (2015). “A General Attraction Model and Sales-Based Linear Program for Network Revenue Management Under Customer Choice”. In: *Operations Research* 63.1, pp. 212–232.

Gallego, Guillermo and Garrett van Ryzin (1997). “A Multiproduct Dynamic Pricing Problem and Its Applications to Network Yield Management”. In: *Operations Research* 45.1, pp. 24–41.

Gallego, Guillermo and Garrett van Ryzin (1994). “Optimal Dynamic Pricing of Inventories with Stochastic Demand over Finite Horizons”. In: *Management Science* 40.8, pp. 999–1020.

Gallego, Guillermo and Garrett Van Ryzin (1994). “Optimal dynamic pricing of inventories with stochastic demand over finite horizons”. In: *Management Science* 40.8, pp. 999–1020.

Gamarnik, David, Assaf Zeevi, et al. (2006). “Validity of heavy traffic steady-state approximations in generalized Jackson networks”. In: *The Annals of Applied Probability* 16.1, pp. 56–90.

Greene, William H. and David A. Hensher (2003). “A latent class model for discrete choice analysis: contrasts with mixed logit”. In: *Transportation Research Part B: Methodological* 37.8, pp. 681–698.

Gurvich, Itai and Amy Ward (2014). “On the dynamic control of matching queues”. In: *Stochastic Systems* 4.2, pp. 479–523.

Gurvich, Itay and Ward Whitt (2009). “Scheduling flexible servers with convex delay costs in many-server service systems”. In: *Manufacturing & Service Operations Management* 11.2, pp. 237–253.

Hajek, Bruce (2006). *Lecture Notes for ECE 467 Communication Network Analysis*.

Harrison, J Michael (2013). *Brownian models of performance and control*. Cambridge University Press.

Hazan, Elad (2016). “Introduction to Online Convex Optimization”. In: *Foundations and Trends® in Optimization* 2.3-4, pp. 157–325.

Hensher, D. and W. Greene (2002). “The Mixed Logit model: The state of practice”. In: *Transportation* 30, pp. 133–176.

Hu, Bin, Ming Hu, and Han Zhu (2019). “Surge pricing and two-sided temporal responses in ride-hailing”. In: *Available at SSRN 3278023*.

Hu, Ming and Yun Zhou (2018). “Dynamic type matching”. Rotman School of Management Working Paper No. 2592622.

Huber, Joel, John W. Payne, and Christopher Puto (1982). “Adding Asymmetrically Dominated Alternatives: Violations of Regularity and the Similarity Hypothesis”. In: *Journal of Consumer Research* 9.1, pp. 90–98.

Jagabathula, Srikanth and Paat Rusmevichientong (2019). “The Limit of Rationality in Choice Modeling: Formulation, Computation, and Implications”. In: *Management Science* 65.5, pp. 2196–2215.

Jaggi, Martin (2013). “Revisiting Frank-Wolfe: Projection-Free Sparse Convex Optimization”. In: *Proceedings of the 30th International Conference on Machine Learning*. Ed. by Sanjoy Dasgupta and David McAllester. Vol. 28. Proceedings of Machine Learning Research 1. Atlanta, Georgia, USA: PMLR, pp. 427–435.

Jasin, Stefanus (2014). “Reoptimization and Self-Adjusting Price Control for Network Revenue Management”. In: *Operations Research* 62.5, pp. 1168–1178.

Jasin, Stefanus (2015). “Performance of an LP-Based Control for Revenue Management with Unknown Demand Parameters”. In: *Operations Research* 63.4, pp. 909–915.

Jasin, Stefanus and Sunil Kumar (2012). “A Re-Solving Heuristic with Bounded Revenue Loss for Network Revenue Management with Customer Choice”. In: *Mathematics of Operations Research* 37.2, pp. 313–345.

Jasin, Stefanus and Sunil Kumar (2013). “Analysis of Deterministic LP-Based Booking Limit and Bid Price Controls for Revenue Management”. In: *Operations Research* 61.6, pp. 1312–1320.

Kanoria, Yash and Pengyu Qian (2019). “Near Optimal Control of a Ride-Hailing Platform via Mirror Backpressure”. In: *arXiv preprint arXiv:1903.02764*.

Kennington, Jeff L. (1978). “A Survey of Linear Cost Multicommodity Network Flows”. In: *Operations Research* 26.2, pp. 209–236.

Kim, Jeunghyun and Ramandeep S Randhawa (2017). “The value of dynamic pricing in large queueing systems”. In: *Operations Research* 66.2, pp. 409–425.

Kleinberg, Robert (2005). “A Multiple-Choice Secretary Algorithm with Applications to Online Auctions”. In: *Proceedings of the Sixteenth Annual ACM-SIAM Symposium on*

Discrete Algorithms. SODA '05. Vancouver, British Columbia: Society for Industrial and Applied Mathematics, 630–631. ISBN: 0898715857.

Kleywegt, Anton J. and Jason D. Papastavrou (1998). "The Dynamic and Stochastic Knapsack Problem". In: *Operations Research* 46.1, pp. 17–35.

Krile, Srećko (2004). "Application of The Minimum Cost Flow Problem in Container Shipping". In: *Proceedings. Elmar-2004. 46th International Symposium on Electronics in Marine*. IEEE, pp. 466–471.

Lange, Daniela Hurtado and Siva Theja Maguluri (2019). "Heavy-traffic Analysis of the Generalized Switch under Multidimensional State Space Collapse". In: *ACM SIGMETRICS Performance Evaluation Review* 47.2, pp. 36–38.

Levitin, Evgeny S. and Boris T. Polyak (1966). "Constrained Minimization Methods". In: *USSR Computational Mathematics and Mathematical Physics* 6.5, pp. 1–50.

Lin, Cheng-Chang and Shwu-Chiou Lee (2015). "Zone Pricing for Time-Definite LTL Freight Transportation with Elastic Demand". In: *Computers & Operations Research* 62, pp. 51–60.

Lin, Cheng-Chang, Dung-Ying Lin, and Melanie M. Young (2009). "Price Planning for Time-Definite Less-than-Truckload Freight Services". In: *Transportation Research Part E: Logistics and Transportation Review* 45.4, pp. 525–537.

Liu, Qian and Garrett van Ryzin (2008). "On the Choice-Based Linear Programming Model for Network Revenue Management". In: *Manufacturing & Service Operations Management* 10.2, pp. 288–310.

Low, David W (1974a). "Optimal dynamic pricing policies for an M/M/s queue". In: *Operations Research* 22.3, pp. 545–561.

Low, David W. (1974b). "Optimal pricing for an unbounded queue". In: *IBM Journal of research and Development* 18.4, pp. 290–302.

Luce, R. Duncan (1959). "On the Possible Psychophysical Laws." In: *Psychological Review* 66.2, pp. 81–95.

Maas, Andrew L., Awni Y. Hannun, and Andrew Y. Ng (2013). "Rectifier nonlinearities improve neural network acoustic models". In: *in ICML Workshop on Deep Learning for Audio, Speech and Language Processing*.

Maglaras, Constantinos and Joern Meissner (2006). "Dynamic Pricing Strategies for Multiproduct Revenue Management Problems". In: *Manufacturing & Service Operations Management* 8.2, pp. 136–148.

Maguluri, Siva Theja and R Srikant (2015). “Queue length behavior in a switch under the maxweight algorithm”. In: *arXiv preprint arXiv:1503.05872*.

— (2016). “Heavy traffic queue length behavior in a switch under the MaxWeight algorithm”. In: *Stochastic Systems* 6.1, pp. 211–250.

McFadden, Daniel (1974). “Conditional logit analysis of qualitative choice behavior”. In: *Frontiers in Econometrics*, pp. 105–142.

McFadden, Daniel and Kenneth Train (2000). “Mixed MNL Models for Discrete Response”. In: *Journal of Applied Econometrics* 15.5, pp. 447–470.

Mitra, Debasis, K.G. Ramakrishnan, and Qiong Wang (2001). “Combined Economic Modeling and Traffic Engineering: Joint Optimization of Pricing and Routing in Multi-Service Networks”. In: *Teletraffic Engineering in the Internet Era*. Ed. by Jorge Moreira de Souza, Nelson L.S. da Fonseca, and Edmundo A. de Souza e Silva. Vol. 4. Teletraffic Science and Engineering. Elsevier, pp. 73–85.

Nguyen, Lam M and Alexander L Stolyar (2018). “A queueing system with on-demand servers: local stability of fluid limits”. In: *Queueing Systems* 89.3-4, pp. 243–268.

Nijs, Vincent R., Shuba Srinivasan, and Koen Pauwels (2007). “Retail-Price Drivers and Retailer Profits”. In: *Marketing Science* 26.4, pp. 473–487.

Osogami, Takayuki and Makoto Otsuka (2014). “Restricted Boltzmann machines modeling human choice”. In: *Advances in Neural Information Processing Systems*. Ed. by Z. Ghahramani, M. Welling, C. Cortes, N. Lawrence, and K. Q. Weinberger. Vol. 27. Curran Associates, Inc.

Ouorou, Adamou, Philippe Mahey, and J.-Ph. Vial (2000). “A Survey of Algorithms for Convex Multicommodity Flow Problems”. In: *Management Science* 46.1, pp. 126–147.

Özkan, Erhun and Amy R Ward (2020). “Dynamic matching for real-time ride sharing”. In: *Stochastic Systems* 10.1, pp. 29–70.

Paschalidis, I Ch and John N Tsitsiklis (2000). “Congestion-dependent pricing of network services”. In: *IEEE/ACM Transactions on networking* 8.2, pp. 171–184.

Puterman, Martin L. (1994). *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. New York, NY, USA: John Wiley & Sons, Inc.

Reiman, Martin I. and Qiong Wang (2008). “An Asymptotically Optimal Policy for a Quantity-Based Network Revenue Management Problem”. In: *Mathematics of Operations Research* 33.2, pp. 257–282.

Resende, Mauricio G. C. and Panos M. Pardalos, eds. (2006). *Handbook of Optimization in Telecommunications*. Springer US.

Revuz, Daniel and Marc Yor (1999). *Continuous Martingales and Brownian Motion*. Springer Berlin Heidelberg.

Roth, Alvin E, Tayfun Sönmez, and M Utku Ünver (2007). "Efficient kidney exchange: Coincidence of wants in markets with compatibility-based preferences". In: *American Economic Review* 97.3, pp. 828–851.

Schrijver, Alexander (1998). *Theory of linear and integer programming*. John Wiley & Sons.

Secomandi, Nicola (2008). "An Analysis of the Control-Algorithm Re-solving Issue in Inventory and Revenue Management". In: *Manufacturing & Service Operations Management* 10.3, pp. 468–483.

Sharkey, Thomas C. (2011). "Network Flow Problems with Pricing Decisions". In: *Optimization Letters* 5.1, pp. 71–83.

Shevtsova, Irina (2011). "On the Absolute Constants in the Berry-Esseen Type Inequalities for Identically Distributed Summands". In: *arXiv preprint arXiv:1111.6554*.

Shi, Cong, Yehua Wei, and Yuan Zhong (2019). "Process flexibility for multiperiod production systems". In: *Operations Research* 67.5, pp. 1300–1320.

Sivaraman, Vibhaalakshmi, Shaileshh Bojja Venkatakrishnan, Kathleen Ruan, Parimarjan Negi, Lei Yang, Radhika Mittal, Giulia Fanti, and Mohammad Alizadeh (2020). "High Throughput Cryptocurrency Routing in Payment Channel Networks". In: *17th USENIX Symposium on Networked Systems Design and Implementation (NSDI 20)*, pp. 777–796.

Srikant, R. and Lei Ying (2014). *Communication Networks: An Optimization, Control and Stochastic Networks Perspective*. New York, NY, USA: Cambridge University Press.

Stolyar, Alexander L et al. (2004). "Maxweight scheduling in a generalized switch: State space collapse and workload minimization in heavy traffic". In: *The Annals of Applied Probability* 14.1, pp. 1–53.

Talluri, Kalyan and Garrett van Ryzin (1998). "An Analysis of Bid-Price Controls for Network Revenue Management". In: *Management Science* 44.11-part-1, pp. 1577–1593.

Talluri, Kalyan and Garrett van Ryzin (2004). "Revenue Management Under a General Discrete Choice Model of Consumer Behavior". In: *Management Science* 50.1, pp. 15–33.

Talluri, Kalyan T and Garrett J Van Ryzin (2004). *The Theory and Practice of Revenue Management*. Springer US.

Tan, Bo and Rayadurgam Srikant (2012). "Online advertisement, optimization and stochastic networks". In: *IEEE Transactions on Automatic Control* 57.11, pp. 2854–2868.

Tassiulas, L. and A. Ephremides (1992). "Stability properties of constrained queueing systems and scheduling policies for maximum throughput in multihop radio networks". In: *IEEE Transactions on Automatic Control* 37.12, pp. 1936–1948.

Tversky, Amos, Paul Slovic, and Daniel Kahneman (1990). "The Causes of Preference Reversal". In: *The American Economic Review* 80.1, pp. 204–217.

Vaidyanathan, Balachandran, Krishna C. Jha, and Ravindra K. Ahuja (2007). "Multicommodity Network Flow Approach to the Railroad Crew-Scheduling Problem". In: *IBM Journal of Research and Development* 51.3.4, pp. 325–344.

Vera, Alberto and Siddhartha Banerjee (2021). "The Bayesian Prophet: A Low-Regret Framework for Online Decision Making". In: *Management Science* 67.3, pp. 1368–1391.

Wieberneit, Nicole (2008). "Service Network Design for Freight Transportation: A Review". In: *OR Spectrum* 30.1, pp. 77–112.

Williams, Ruth J (1998). "Diffusion approximations for open multiclass queueing networks: sufficient conditions involving state space collapse". In: *Queueing systems* 30.1-2, pp. 27–88.

Williamson, Elizabeth Louise (1992). "Airline Network Seat Inventory Control: Methodologies and Revenue Impacts". PhD thesis. Massachusetts Institute of Technology.

Wu, Huasen, R. Srikant, Xin Liu, and Chong Jiang (2015). "Algorithms with Logarithmic or Sublinear Regret for Constrained Contextual Bandits". In: *Proceedings of the 28th International Conference on Neural Information Processing Systems - Volume 1*. NIPS'15. Montreal, Canada: MIT Press, 433–441.

Yan, Chiwei, Helin Zhu, Nikita Korolko, and Dawn Woodard (2019). "Dynamic pricing and matching in ride-hailing platforms". In: *Naval Research Logistics (NRL)*. (forthcoming).

Zinkevich, Martin (2003). "Online Convex Programming and Generalized Infinitesimal Gradient Ascent". In: *Proceedings of the Twentieth International Conference on International Conference on Machine Learning*. ICML'03. Washington, DC, USA: AAAI Press, 928–935. ISBN: 1577351894.

Zohar, Ety, Avishai Mandelbaum, and Nahum Shimkin (2002). “Adaptive behavior of impatient customers in tele-queues: Theory and empirical support”. In: *Management Science* 48.4, pp. 566–583.